Advances in Intelligent Systems and Computing 476

Helena Lindgren
Juan F. De Paz
Paulo Novais
Antonio Fernández-Caballero
Hyun Yoe
Andres Jiménez Ramírez
Gabriel Villarrubia *Editors*

Ambient Intelligence-Software and Applications – 7th International Symposium on Ambient Intelligence (ISAmI 2016)

Springer

Advances in Intelligent Systems and Computing

Volume 476

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Chairman
Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in
Members
Rafael Bello, Universidad Central "Marta Abreu" de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu
Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchado@usal.es
Hani Hagras, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk
László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu
Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu
Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlin@mail.nctu.edu.tw
Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au
Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org
Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br
Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl
Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

More information about this series at http://www.springer.com/series/11156

Helena Lindgren · Juan F. De Paz
Paulo Novais · Antonio Fernández-Caballero
Hyun Yoe · Andres Jiménez Ramírez
Gabriel Villarrubia
Editors

Ambient Intelligence - Software and Applications – 7th International Symposium on Ambient Intelligence (ISAmI 2016)

Springer

Editors
Helena Lindgren
Plan 4 MIT-huset C435
University of Umea
Umea
Sweden

Juan F. De Paz
Dept de Info y Automática
Universidad de Salamanca
Salamanca
Spain

Paulo Novais
Dept de Informatica, Campus of Gualtar
Universidade do Minho
Braga
Portugal

Antonio Fernández-Caballero
Departamento de Sistemas Informáticos
University of Castilla-La Mancha
Albacete
Spain

Hyun Yoe
Department of Information and Communication
Sunchon National University
540-742
Korea (Republic of)

Andres Jiménez Ramírez
ETS Ingeniería Informática
University of Sevilla
Sevilla
Spain

Gabriel Villarrubia
Departamento de Informática y Automática
Universidad de Salamanca
Salamanca
Spain

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-40113-3 ISBN 978-3-319-40114-0 (eBook)
DOI 10.1007/978-3-319-40114-0

Library of Congress Control Number: 2016940894

© Springer International Publishing Switzerland 2016
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

This volume contains the proceedings of the 7th International Symposium on Ambient Intelligence (ISAmI 2016). The Symposium was held in Seville, Spain on June 1st–3rd at the University of Seville.

ISAmI has been running annually and aiming to bring together researchers from various disciplines that constitute the scientific field of Ambient Intelligence to present and discuss the latest results, new ideas, projects and lessons learned, namely in terms of software and applications, and aims to bring together researchers from various disciplines that are interested in all aspects of this area.

Ambient Intelligence is a recent paradigm emerging from Artificial Intelligence, where computers are used as proactive tools assisting people with their day-to-day activities, making everyone's life more comfortable.

After a careful review, 25 papers from 12 different countries were selected to be presented in ISAmI 2016 at the conference and published in the proceedings. Each paper has been reviewed by, at least, three different reviewers, from an international committee composed of 129 members from 24 countries.

Acknowledgments
A special thanks to the editors of the workshops (SE4GT) Smart Environments for Gerontechnology and e-Therapy.

This event was partially supported by the project, EKRUCAmI: Europe-Korea Research on Ubiquitous Computing and Ambient Intelligence. The project has been supported by European Commission (Seventh Framework Programme for Research and Technological Development. FP7-PEOPLE-2012-IRSES. Marie Curie Action "International Research Staff Exchange Scheme").

We want to thank all the sponsors of ISAmI'16: IEEE Sección España, CNRS, AFIA, AEPIA, APPIA, AI*IA, and Junta de Castilla y León.

ISAmI would not have been possible without an active Program Committee. We would like to thanks all the members for their time and useful comments and recommendations.

We would like also to thank all the contributing authors and the Local Organizing Committee for their hard and highly valuable work.

Your work was essential to the success of ISAmI 2016.

June 2016

Helena Lindgren
Juan F. De Paz
Paulo Novais
Antonio Fernández-Caballero
Hyun Yoe
Andrés Jimenez-Ramírez
Gabriel Villarrubia

Organization

Program Committee Chairs

Juan F. De Paz Santana (Chair)	University of Salamanca, Spain
Helena Lindgren	Umeå University, Sweden

Program Committee

Adina Magda Florea	University Politehnica of Bucharest, AI-MAS Laboratory
Ahlem Melouah	University of Annaba
Alberto Lopez	University of Salamanca
Alejandro Catala	University of Castilla-La Mancha
Alexandru Sorici	Faculty of Automatic Control and Computers, University “Politehnica” of Bucharest
Alfonso Gonzalez Briones	BISITE Research Group
Ali Salah	Boğaziçi University
Álvar Legaz-Aparicio	Universidad Politécnica de Cartagena
Alvaro Lozano	USAL
Álvaro Castro González	Universidad Carlos III de Madrid
Ana Coelho	University of Minho
Ana Almeida	GECAD-ISEP-IPP
Ana Alves	Centre for Informatics and Systems of University of Coimbra
Anastasios Karakostas	Department of Informatics, Aristotle University of Thessaloniki, Greece
Antonio Fernandez Caballero	Universidad de Castilla-La Mancha

Antonio Meireles	ISEP
Antonio Pereira	Escola Superior de Tecnologia e Gestão do IPLeiria
António Abelha	Universidade do Minho
António Pinto	INESC Porto
Arturo C. Rodríguez	Universidad De Castilla-La Mancha
Bogdan Kwolek	AGH University of Science and Technology
Celestino Goncalves	Instituto Politecnico da Guarda
Cesar Analide	University of Minho
Clarence W. de Silva	UBC Mechanical Engineering
Daniel Hernandez	University of Salamanca
Danilo Avola	Dipartimento Informatica - Sapienza - Università di Roma
David Davide	Universidade do Minho
David Griol	Universidad Carlos III de Madrid, Spain
Davide Carneiro	Universidade do Minho
Diego Martínez-Plasencia	University of Sussex
Dong Feng Liu	Thorlabs
Elena Navarro	University of Castilla-La Mancha
Emilia I. Barakova	Eindhoven University of Technology
Enzo Pasquale Scilingo	University of Pisa
Fábio Silva	Universidade do Minho
Fernando De La Prieta	University of Salamanca
Fernando Garcia-Sanjuan	Universitat Politècnica de València
Fernando Silva	Dept. of Informatics Engineering; School of Technology and Management; Polytechnic Institute of Leiria, Portugal
Fernando De La Torre	Carnegie Mellon University
Ficco Massimo	Second University of Naples (SUN)
Filipe Portela	Algoritmi Research Centre, University of Minho
Filippo Cavallo	Scuola Superiore Sant : Anna
Florentino Fdez-Riverola	University of Vigo
Francisco J.Perales	Universitat de les Illes Balears
Gabriel Villarrubia	University of Salamanca
Georgios Meditskos	Information Technologies Institute, Centre for Research & Technology : Hellas
Goreti Marreiros	ISEP/IPP-GECAD
Guillaume Lopez	Aoyama Gakuin University, College of Science and Technology
Habib M. Fardoun	King Abdulaziz University
Hans W. Guesgen	Massey University
Héctor Gómez	Universidad Técnica Particular de Loja
Helena Lindgren	Department of Computing Science, Umeå University

Ichiro Satoh	National Institute of Informatics
Igor Kotenko	St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS)
Illias Maglogiannis	University of Piraeus Athens
Jaderick Pabico	University of the Philippines Los Banos
Jason Jung	Chung-Ang University
Javier Prieto	University of Salamanca
Javier Jaén	Universitat Politècnica de València
Javier Carbo	Univ. Carlos III of Madrid
Javier Jaen	Polytechnic University of Valencia
Javier Bajo	Universidad Politécnica de Madrid
Jean-Paul Jamont	LCIS, Université de Grenoble
Jesus Favela	CICESE
Jordi Palacin	Universitat de Lleida
Jorge Revuelta	BISITE Research Group
Jose Antonio Castellanos	University of Salamanca
Jose Neves	Universidade do Minho
José Machado	University of Minho, Computer Science and Technology Centre
José Barbosa	Universidad Técnica Particular de Loja
José Manuel Pastor	Universidad de Castilla-La Mancha
José Miguel Latorre	Universidad de Castilla-La Mancha
José Manuel Molina	Universidad Carlos III
Jose Carlos Castillo Montoya	Universidad Carlos III de Madrid
Jose M. Molina	Universidad Carlos III de Madrid
Jose Manuel Pastor Garcia	Universidad de Castilla-La Mancha
José Pascual Molina Massó	Universidad de Castilla-La Mancha
Joyca Lacroix	Philips Research
Juan José Pantrigo	Universidad Rey Juan Carlos
Juan Pavón	Universidad Complutense de Madrid
Juan Ramos	BISITE Research Group
Juan Francisco De Paz Santana	University of Salamanca
Julie Doyle	Centre for Affective Solutions for Ambient Living Awareness
Junzhong Gu	East China Normal University
Koen van Dam	Imperial College London
Lars Braubach	University of Hamburg
Latif Ladid	UL & IPv6 Forum - IOT subC
Lawrence Wai-Choong Wong	National University of Singapore
Lino Figueiredo	ISEP
Lourdes Borrajo	University of Vigo
Luis Orozco-Barbosa	Universidad de Castilla-La Mancha
Luis Seabra Lopes	IEETA - DETI, UA
Manuel Filipe Santos	University of Minho

Marco Ceccarelli	Università degli studi di Cassino e del Lazio
Maria Navarro	BISITE
Mark Hoogendoorn	Vrije Universiteit Amsterdam
Miguel Angel Patricio	Universidad Carlos III
Miguel Angel Salichs	Univ. Carlos III
Nadia Bianchi-Berthouze	University College London
Nayat Sanchez-Pi	Instituto de Lógica, Filosofia e Teoria da Ciência
Nuno Garcia	Universidade Lusofona de Humanidades e Tecnologias
Pablo Chamoso	BISITE
Pascual González	Castilla-La Mancha University
Patricio Domingues	ESTG - Leiria
Paulo Novais	University of Minho
Pawel Pawlewski	Poznan University of Technology
Phan Cong-Vinh	Nguyen Tat Thanh University
Radu-Daniel Vatavu	University Stefan cel Mare of Suceava
Rafael Martínez-Tomás	Universidad Nacional de Educación a Distancia
Rene Meier	Lucerne University of Applied Sciences
Ricardo Anacleto	ISEP/GECAD
Roberto Zangroniz	Universidad de Castilla-La Mancha
Rui José	University of Minho
S. Shyam Sundar	Penn State University, USA & Sungkyunkwan University, Korea
Salvatore Venticinque	Seconda Università di Napoli
Sotiris Nikoletseas	University of Patras and CTI
Sylvie Ratté	École de technologie supérieure
Teresa Romão	DI/FCT/UNL
Thanos Stavropoulos	ITI
Tiancheng Li	Northwestern Polytechnical University
Tibor Bosse	Vrije Universiteit Amsterdam
Tino Lourens	TiviPe
Valérie Camps	University of Toulouse - IRIT, France
Vicente Julian	GTI-IA DSIC UPV
Victor Alves	Universidade do Minho
Víctor Corcoba Magaña	University Carlos III
Wolfgang Minker	University of Ulm
Yiannis Kompatsiaris	CERTH – ITI

Organizing Committee Chair

Paulo Novais	University of Minho, Portugal)
Andrés Jiménez Ramírez	University of Seville, Spain)

Local Organization Committee

Carlos Arevalo Maldonado	University of Sevilla, Spain
Gustavo Aragon Serrano	University of Sevilla, Spain
Irene Barba	University of Sevilla, Spain
Miguel Ángel Barcelona Liédana	Technological Institute of Aragon, Zaragoza, Spain
Juan Manuel Cordero Valle	University of Sevilla, Spain
Francisco José Domínguez Mayo	University of Sevilla, Spain
Juan Pablo Domínguez Mayo	University of Sevilla, Spain
Manuel Domínguez Muñoz	University of Sevilla, Spain
María José Escalona Cuaresma	University of Sevilla, Spain
José Fernández Engo	University of Sevilla, Spain
Laura García Borgoñón	Technological Institute of Aragon, Zaragoza, Spain
Julian Alberto García García	University of Sevilla, Spain
Javier García-Consuegra Angulo	University of Sevilla, Spain
José González Enríquez	University of Sevilla, Spain
Tatiana Guardia Bueno	University of Sevilla, Spain
Andrés Jiménez Ramírez	University of Sevilla, Spain
Javier Jesús Gutierrez Rodriguez	University of Sevilla, Spain
Manuel Mejías Risoto	University of Sevilla, Spain
Laura Polinario	University of Sevilla, Spain
José Ponce Gonzalez	University of Sevilla, Spain
Francisco José Ramírez López	University of Sevilla, Spain
Isabel Ramos Román	University of Sevilla, Spain
Jorge Sedeño López	University of Sevilla, Spain
Nicolás Sánchez Gómez	University of Sevilla, Spain
Juan Miguel Sánchez Begines	University of Sevilla, Spain
Eva-Maria Schön	University of Sevilla, Spain
Jesús Torres Valderrama	University of Sevilla, Spain
Carmelo Del Valle Sevillano	University of Sevilla, Spain
Antonio Vázquez Carreño	University of Sevilla, Spain
Carlos Torrecilla Salinas	University of Sevilla, Spain
Ainara Aguirre Narros	University of Sevilla, Spain
Juan Antonio Alvarez García	University of Sevilla, Spain
Cesar Analide	Univerity of Minho, Portugal

Workshops

SE4GT**: Smart Environments for Gerontechnology and e-Therapy**

Organizing Committee

Antonio Fernández-Caballero	University of Castilla-La Mancha, Spain
Pascual González	University of Castilla-La Mancha, Spain
José Manuel Ferrández	Universidad Politécnica de Cartagena, Spain
Elena Navarro	University of Castilla-La Mancha, Spain

ISAMI 2016 Sponsors

Contents

A Benchmark Dataset for Human Activity Recognition and Ambient Assisted Living

Giuseppe Amato, Davide Bacciu, Stefano Chessa, Mauro Dragone, Claudio Gallicchio, Claudio Gennaro, Hector Lozano, Alessio Micheli, Gregory M.P. O'Hare, Arantxa Renteria and Claudio Vairo

Abstract We present a data benchmark for the assessment of human activity recognition solutions, collected as part of the EU FP7 RUBICON project, and available to the scientific community. The dataset provides fully annotated data pertaining to numerous user activities and comprises synchronized data streams collected from a highly sensor-rich home environment. A baseline activity recognition performance obtained through an Echo State Network approach is provided along with the dataset.

Keywords Ambient assisted living · Human Activity Recognition · Datasets

1 Introduction

Several works in ambient assisted living (AAL) and smart environments areas have harnessed machine learning techniques to merge and interpret the information gathered by multiple noisy sensors [1]. The goal of Human Activity Recognition

G. Amato · S. Chessa(✉) · C. Gennaro · C. Vairo
CNR-ISTI, via Moruzzi 1, 56124 Pisa, Italy
e-mail: stefano.chessa@unipi.it

D. Bacciu · S. Chessa · C. Gallicchio · A. Micheli
Department of Computer Science, University of Pisa, Largo Pontecorvo 3, Pisa, Italy

M. Dragone
Sensors, Signals and Systems, Heriot-Watt University, Edinburgh, UK

H. Lozano · A. Renteria
Tecnalia, Bilbao, Spain

G.M.P. O'Hare
University College Dublin, Dublin, Ireland

This work has been supported by the EU FP7 RUBICON project (contract no. 269914).
© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_1

(HAR) is to determine the state of the users and possibly predict their behaviour, in order to deliver context-aware, proactive and customized assistance [2]. HAR can also be used to detect of anomalies and deviations in normal users' behaviour, which may be symptoms of problems worthy of notification to carers or to the users directly [3, 4].

A common approach to developing HAR systems is to use sensor logs gathered during real-world experiments comprising a user performing a number of activity patterns [5]. These data sets contains the sensor data and the annotations (usually recorded by human observers) describing the activities performed by the users. Such data typically serve to train machine learning models to predict the occurrence of the human activities from the sensor data, and to assess their predictive and generalization performance.

Several works published case studies comprising the prediction of human daily living activities (ADL) from data gathered in sensorised environments [4, 6 – 9]. However, only a few of them include datasets collected from real systems deployed in real homes with real users (such as [9]), and were made freely available to the scientific community. In addition, available datasets are mostly limited to basic monitoring services and homogeneous settings (e.g. reporting only data from binary sensors). On the other hand, the cost and the technical problems due to data annotation, to the design and of the experimental procedures, and to the data collection software, make these datasets an extremely valuable tool for developing integrated smart environments.

The EU project RUBICON [4, 21] develops cognitive robotic ecologies, i.e. systems made out of wireless sensors, actuators and robots, each equipped with perception, attention, memory, action, learning, and planning capabilities. By following a multi-agent approach, these 'robotic devices' together can provide complex services, by leveraging their ability to acquire and apply knowledge about their physical settings, and to coordinate the actions between themselves. In order to promote the development of cognitive robotic ecologies and, more generally, of AAL and smart home environments, this paper describes the dataset that was recently collected and used to assess state-of-the-art HAR solutions developed as part of the RUBICON system, that we now share with the research community. The complete dataset contains fully labelled data pertaining to numerous user's activities. It reports 139 synchronized data streams, obtained by heterogeneous sensors (both embedded and wearable, binary and analogue, wired and wireless networks), sound recognition components and effectors and robots.

In addition, in order to provide a baseline for the predictive performance on the dataset, the paper also briefly discusses the results achieved by the Echo State Network (ESN) model on this benchmark dataset. The ESN neural network model [10] is a state-of-the-art approach for efficient time-series processing, successfully used in many application areas (e.g. chaotic time series prediction [10, 11], robot localization [12], control systems [13], AAL [14, 15, 20]), HAR [16], etc.), although achieved on other datasets. The use of ESNs in this context is motivated by the fact that they are particularly suitable for dealing with tasks in which the history of past input signals plays a key role in the determination of the ongoing activity, performing sensor data fusion and noise-robust processing of both discrete and real data from heterogeneous sources. As the model presentation is out of the scope of this work, for details on the ESN approach, its architecture and properties, the reader is referred to the literature [10,11].

2 The Test-Bed

The test-bed used to collect the dataset is the TECNALIA HomeLab – a fully functional apartment (shown in Figure 1) of over $45m^2$ with a bed-room, livingroom, kitchen area, bathroom, and a corridor. It is equipped with embedded sensors, effectors, and with two mobile robots. A hybrid (wireless/wired) communication infrastructure [17] is used to read data gathered from the wearable and environmental sensors. The environmental sensors include switches signaling opening/closing of drawers/doors and occupancy sensors detecting the room of the user. The communication infrastructure also connects effectors, such as lights, blinds, door locks, robots and appliances to allow sending commands and receiving their status. In particular, this allows also to detect when the user interacts with the effectors (for example she manually switches on or off the TV and lights, or when she requests a service from one of the robots).

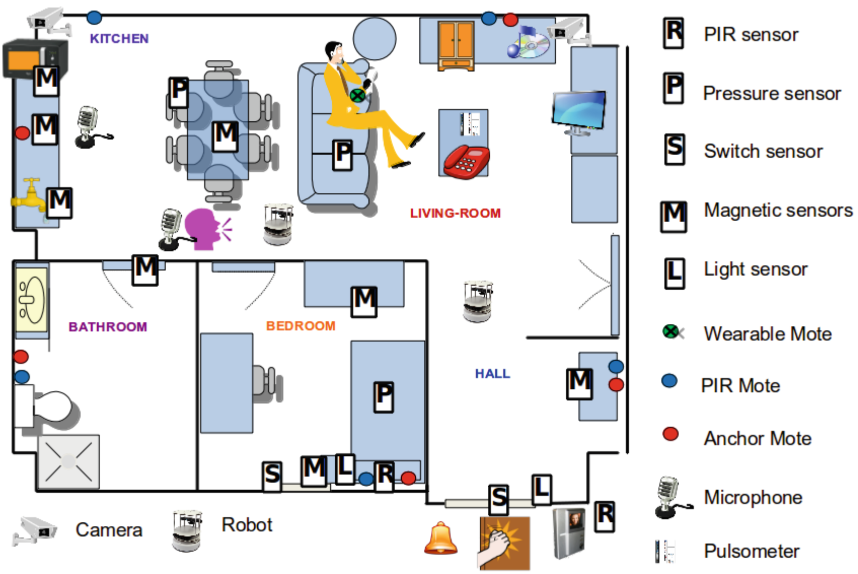


Fig. 1 Location of the sensors and actuators in the HomeLab

In such a context, the learning abilities of the RUBICON system are used to identify and react to user needs, activities, and preferences, and to support the adaptive and flexible operations of a range of AAL services. For instance, RUBICON can learn to activate automatically appliances and/or robotic services in relation to the user's actions. For example, RUBICON can learn to close the blinds when the user is sleeping, clean the floor only after the user has had her meal, remind the user to measure her pulse and blood pressure, or to fetch the pulsometer and pressure monitor if the user is indisposed. The following is the list of devices installed in the HomeLab:

- An IEEE 802.15.4 wireless sensor network, with 11 Advanticsys[1] sensor motes compliant with the TelosB architecture [18]. This includes: 5 CM3000 motes equipped with SE1000 sensor boards (providing readings from a passive infrared (PIR) occupancy sensor); 5 CM3000 anchor motes transmitting beacon packets at regular intervals of 500 ms; 1 CM3000 mote installed on a bracelet worn by the user equipped with an EM1000 sensor board providing information from a 2-axis accelerometer.
- A Konnex (KNX home automation system[2]) and a proprietary domotic network, with a total of 30 sensing and effector devices, including: 2 additional PIRs, 3 pressure sensors, 14 magnetic sensors, 8 general purpose input/output modules, 1 microphone, 1 bluetooth-enabled pulsometer and blood pressure monitor.
- 2 Turtlebot robots[3], equipped with R-GBD cameras.

In addition, 2 ceiling cameras are installed at opposite sides of the HomeLab. These cameras have not been used as input sources for our AAL solutions, but only to provide a record of all the experiments in the HomeLab.

3 The Experimental Procedure

Scripts – Our experimental procedure is coherent with and extends those pursued in similar initiatives [5]. We log datasets from typical household scenarios consisting in temporally unfolding situations made out of both complex (e.g. prepare meal) and atomic activities (e.g. reach for bread cutter, use bread cutter, use toaster). Naturalistic execution of each scenario is achieved by relying on hierarchical scripts and leaving free interpretation to one volunteer that was recruited to perform a set of trials in the test-bed. Each script describes a sequence of activities (e.g. washing dishes, cleaning, resting...) in terms of high level actions to be performed by the user (e.g. use the scrubber/the detergent, sit on the chair, clean table, put back scrubber. ..), leaving space for a natural execution of the task. For instance, each script allows variations on the way the script is performed, such as changing, to some extent, the sequence of actions involved (e.g. open one drawer before the other), or using different paths (e.g. left or right of the sofa) to leave the Homelab from the kitchen.

In addition, scripts contains instructions describing the operation of one Turtlebot robot. These instructions are executed by a Wizard of Oz (WoZ) operator, which remotely controls the robot by using the Rviz tool included in the robotic operating system (ROS) [19]. Rviz also allows the user can 'ask' for the activation of specific robotic services (for instance, the user can request the robot to vacuum clean the

---

[1] http://www.advanticsys.com
[2] http://www.knx.org
[3] http://turtlebot.com

floor of the kitchen), without any speech recognition or graphical user interface, thus allowing us to shield the data collection activities from issues concerning user interfaces.

Logging Infrastructure – A number of precautions to follow to enable logging synchronized data from highly heterogeneous networks, with proprietary, non-interoperable protocols (each characterized by their own latency, and packet loss rate) are given in [5]. We overcome these problems by relying on a purposefully designed logging application integrated with our communication framework. Following the suggestions in [5], our logging application acts as a central monitoring station used to start/stop each experiment, and to provide synchronization at protocol level. To begin with, UDP broadcast on a dedicated LAN relays the start/stop instructions from the monitoring station to all the other elements of our logging framework, respectively:

- KNX: The readings from the sensors and actuators in this network are logged into a relational database using a proprietary software solution.
- Robot: The activities of the Turtlebot are logged in ROS bags on each robot, and transferred to a central repository afterwards.
- Cameras: The videos recorded by the ceiling cameras capture all events occurred during the experimentation.
- WSN: The data acquired by the WSN sensors are collected by a base station where they are stored in a relational database along with their timestamps.

The relational database in the base station of the WSN is organized in two tables: one contains the experiments performed (Experiments entity) and one contains all the data recorded by the sensors in each experiment (DataSample entity). Each experiment is associated with an ID, called ExpID, which is a foreign key in the DataSample table and which provides the association between the two tables. The latter precaution is taken to avoid losing some of the sensor information because of packet losses during the online operations of the WSN. In fact, the actual run-time computation of the RUBICON distributed learning sub-system is performed by learning modules embedded on the sensor nodes and only the external event outputs are delivered to the decisional and executive components of our system [14]. Such an arrangement enables us to avoid transmitting all the sensor data collected by the WSN to a centralized learning module – a solution that would quickly deplete the battery of the motes and likely cause an unmanageable level of packet losses.

Data Annotation – During the data collection campaign, a human operator annotates, by means of a mobile phone application, the user activities and locations to produce the ground truth description of the events stored in the dataset. Furthermore, for each trial, the exact configuration of the system, in terms of sensor number, actual position, types of sensors, sampling rate/type of pre-processing (if any), IDs of devices, version and configuration of the software is also recorded and stored as meta-data.

4 The Dataset

We gathered real-world data for the AAL test-bed in the HomeLab facility presented in Section 2, by collecting raw data from heterogeneous sources at the frequency of 2 Hz (2 data samples per second) while an actor was performing a set of daily-life activities. The only exception to the collection procedure was made for the microphones. In this case, rather than logging the raw sound signal, we recorded only the recognition confidence associated to a list of known sounds (door bell, water pouring, microwave bell, ...) produced by a pre-processing module. In total, the recording of each experiment consists of 139 data streams (66 from the 802.15.4 WSN, and 73 from the KNX).

In order to test the ability of our system to detect users' activities and routines, and learn to assist the users in their daily living, we have run a series of experiments by considering nine different user activities, namely Eating, Set Table, User Cleaning, Washing Dishes, Prepare Food, Exercising, Relaxing, Sleeping and User Arrives. Five environment locations were also considered for the purpose of user localization prediction, i.e. Outside, Hall, Living Room, Kitchen, Bedroom. Each sequence in the dataset contains the measurements from sensors and actuators and the (manually labelled) ground-truth corresponding to a number of activities consecutively executed by the user, for a total of 83 sequences. In addition, some of these activities involved robot services. Specifically, users were instructed to switch on the TV while eating, or to ask the robot to vacuum clean the kitchen floor after they had finished their meal.

Table 1 Per-activity accuracy and standard deviations achieved by ESN models.

Activity/localization	**Per-activity accuracy**
Eating	91.97% (±0.09)
Set Table	90.31% (±0.10)
User Cleaning	84.25% (±0.22)
Washing Dishes	83.61% (±0.02)
Prepare Food	86.37% (±0.06)
Exercising	94.36% (±2.53)
Relaxing	100% (±0.00)
Sleeping	94.8% (±0.05)
User Arrives	98.81% (±0.02)
Loc. Outside	94.21% (±0.98)
Loc. Hall	85.78% (±3.58)
Loc. Living room	88.54% (±1.25)
Loc. Kitchen	87.94% (±1.98)
Loc. Bedroom	97.47% (±0.33)

The resulting dataset, together with a full description of the test-bed, the scripts, and the baseline performance achieved by ESNs on the dataset (including data splitting into training and test, and input wiring from the available set of sensors for each activity) is freely available to the scientific community[4]. The videos captured from the robot's sensors and from the ceiling cameras can also be obtained on request.

5 Results

The dataset was used to train the activity recognition components of RUBICON (which are based on ESN models) in an AAL demonstrator. We present here an extract of the performance of these ESN models on the dataset, restricted to non-video streams. This may also serve as baseline reference performance focusing on HAR tasks.

For every computational task, the predictive performance, in terms of per-class (or per-activity) accuracy of the ESN prediction, was evaluated on a separate external test set (≈25% of the available data), after a process of hold-out model selection on a validation set (whose size is ≈25% of the training set size). Training was conducted separately for each computational task (activity/location), selecting a specific sub-set of input sources, using the data pertaining to the specific activity (or location) as positive samples, and the data pertaining to other activities (or locations) as negative samples. For each computational task considered, training was performed using a balanced training set. To this aim, for each task, all the positive training samples were used, while the negative samples were uniformly chosen among the remainder of the activities, in such a way that the number of positive and negative samples resulted the same.

The predictive generalization performance obtained on the various tasks considered in the AAL scenario is shown in Table 1, which reports the per-class accuracies on the test set for each activity and localization under consideration. As can be seen from Table 1, the predictive performance is very good for every activity and location, achieving results which are in line with the state of the art in HAR tasks, e.g. as reported in a recent survey [7], although obtained in different experimental settings. The averaged test accuracy is 91.32%(≈0.80), with per-class test accuracies ranging from a minimum of 83.61% to a maximum of 100%, where the lowest accuracies were obtained for activities corresponding to highly noisy input sensors (e.g. Washing Dishes), or corresponding to cases of intrinsic similarities, given the available sensors, among sensorial patterns (e.g. Prepare Food and Set Table). Figure 2 shows a graphical representation of the output of the ESN classification for a relevant subset of activities of interest, where output values above 0 correspond to the recognition of the event (i.e. positively classified sampled), and the absolute output values can be interpreted as the confidence of the classification.

[4] The dataset can be downloaded from
http://fp7rubicon.eu/uploads/HomelabAALdatasetISAMI/HomelabAALdataset.zip

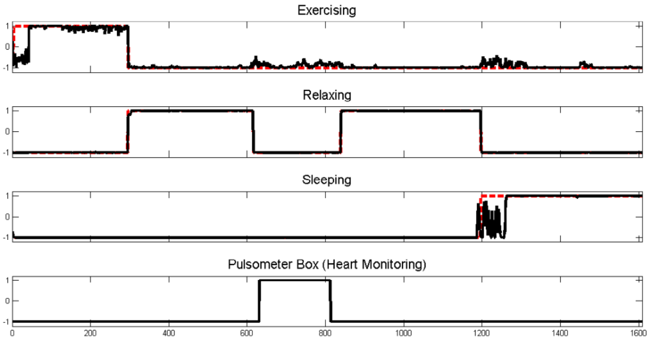


Fig. 2 Output of the ESN classification, corresponding ground-truth for a test set sequence in the benchmark dataset, and input signal from the pulsometer box.

6 Conclusions

The dataset provides an extremely rich collection of sensor information of a heterogeneous nature such that only a limited subset of the available sensor is relevant and predictive for the purpose of recognizing each of the single activities. Therefore, it can also be exploited as a challenging benchmark to evaluate automatic feature selection algorithms for heterogeneous timeseries data, including both binary measurements (e.g. pressure sensor information) and noisy multi-valued data. Furthermore, the dataset can be used to test solutions for the automatic discovery of user's activities and/or automation patterns. To conclude, the authors wish to thank all the other colleagues in the RUBICON consortium: Örebro University (SE); Robotnik Automation SLL (SP); University of Ulster (UK); Pintal LTD (IE); Fondazione Stella Maris (IT).

References

1. Ye, J., Dobson, S., McKeever, S.: Review: Situation identification techniques in pervasive computing: A review. Pervasive Mobile Computing **8**(1), 36–66 (2012)
2. Aztiria, A.: Learning frequent behaviours of the users in intelligent environments. J. Ambient Intelligence and Smart Environments **2**(4), 435–436 (2010)
3. Cook, D.J., Krishnan, N.C., Rashidi, P.: Activity discovery and activity recognition: A new partnership. IEEE Transactions on Cybernetics **43**(3), 820–828 (2013)
4. Rubicon: A Self-learning Robotic Ecology, http://fp7rubicon.eu/, 2011, last referenced on December 20, 2015
5. Roggen, D., et al.: Collecting complex activity datasets in highly rich networked sensor environments. In: IEEE 7th Int. Conf. on Networked Sensing Systems (INSS), pp. 233–240 (2010)
6. Kurz, M., et al.: The opportunity framework and data processing ecosystem for opportunistic activity and context recognition. Int. J. of Sensors, Wireless Communications and Control, Special Issue on Autonomic and Opportunistic Communications (2011)

7. Lara, O.D., Labrador, M.A.: A survey on human activity recognition using wearable sensors. IEEE Communications Surveys & Tutorials **15**(3), 1192–1209 (2013)
8. Bacciu, D., Chessa, S., Gallicchio, C., Lenzi, A., Micheli, A., Pelagatti, S.: A general purpose distributed learning model for robotic ecologies. In: IFAC 10th Int. Symposium on Robot Control (SYROCO 2012), vol. 10, pp. 435–440 (2012)
9. Cook, D.J., Crandall, A.S., Thomas, B.L., Krishnan, N.C.: Casas: A smart home in a box. Computer **46**(7), 62–69 (2013)
10. Jaeger, H., Haas, H.: Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science **304**(5667), 78–80 (2004)
11. Gallicchio, C., Micheli, A.: Architectural and markovian factors of echo state networks. Neural Networks **24**(5), 440–456 (2011)
12. Chessa, S., Gallicchio, C., Guzman, R., Micheli, A.: Robot localization by echo state networks using RSS. In: Recent Advances of Neural Network Models and Applications, ser. Smart Innovation, Systems and Technologies, vol. 26, pp. 147–154. Springer (2014)
13. Joshi, P., Maass, W.: Movement generation with circuits of spiking neurons. Neural Computation **17**(8), 1715–1738 (2005)
14. Bacciu, D., Barsocchi, P., Chessa, S., Gallicchio, C., Micheli, A.: An experimental characterization of reservoir computing in ambient assisted living applications. Neural Computing and Applications **24**(6), 1451–1464 (2014)
15. Bacciu, D., Chessa, S., Gallicchio, C., Micheli, A., Barsocchi, P.: An experimental evaluation of reservoir computation for ambient assisted living. In: Neural Nets and Surroundings, ser. Smart Innovation, Systems and Technologies, vol. 19, pp. 41–50. Springer (2013)
16. Palumbo, F., Barsocchi, P., Gallicchio, C., Chessa, S., Micheli, A.: Multisensor data fusion for activity recognition based on reservoir computing. In: Evaluating AAL Systems Through Competitive Benchmarking, ser. CCIS, vol. 386, pp. 24–35. Springer (2013)
17. Amato, G., Broxvall, M., Chessa, S., Dragone, M., Gennaro, C., Vairo, C.: When wireless sensor networks meet robots. In: 7th Int. Conf. on Systems and Networks Communications (ICSNC 2012). Lisbon, PT, pp. 35–40 (2012)
18. Polastre, J., Szewczyk, R., Culler, D.: Telos: Enabling ultra-low power wireless research. In: 4th Int. Sym. on Information Processing in Sensor Networks (IPSN), UCLA, Los Angeles, California, USA (2005)
19. Quigley, M., et al.: ROS: An open-source Robot Operating System. In: ICRA Workshop on Open Source Software (2009)
20. Dragone, M., et al.: A cognitive robotic ecology approach to self-configuring and evolving AAL systems. Engineering Applications of Artificial Intelligence **45**, 269–280 (2015)
21. Amato, G., et al.: Robotic ubiquitous cognitive ecology for smart homes. Journal of Intelligent & Robotic Systems: Theory and Applications (2015, in press)

An Enhanced Real Space Through Temporally Connecting Real and Virtual Scenes

Fumiko Ishizawa and Tatsuo Nakajima

Abstract The potential power of augmented reality (AR) makes it possible to change the semiotic meaning of the real space, and strongly influence human attitude and behavior for achieving a better lifestyle. Typical AR approaches can simply integrate computer-generated (CG) virtual images on the real space from a spatial aspect. This approach limits the potential power because only some objects in the real space can be enhanced. In contrast, virtual reality (VR) technologies make it possible to create a new fictional, but realistic virtual space by using 3D models of real persons and landscapes. In this paper, we propose a new concept named *Alternative Reality* that is used for building a new type of ambient intelligence experiences. *Alternative Reality* makes it possible to connect the real space with the virtual space temporally and to integrate the two spaces seamlessly, where its virtual space consists of real landscapes and persons not fictional ones to change the semiotic meaning of the real space and strongly influence human attitude and behavior. However, the virtual space does not reflect the events that actually happened in the real space; this means that it may be possible to enhance the real space by showing fictional events among real events. For demonstrating the *Alternative Reality* concept, we have developed a case study named *Fictional Future*, where the virtual space is constructed with fictional events performed by real persons in real landscapes. We conducted a user study of the case study to show the feasibility of the *Alternative Reality* concept. The user study shows that *Alternative Reality* augments the real space effectively and a user still experiences the virtual space as in the real space.

Keywords Virtual Reality · Augmented Reality · Head Mounted Display

1 Introduction

Today, augmented reality (AR) technologies have become popular for developing a variety of ambient intelligence and entertainment services [1]. However, because

---

F. Ishizawa(✉) · T. Nakajima
Department of Computer Science and Engineering, Waseda University,
3-4-1 Okubo, Shinjuku, Tokyo 169-8555, Japan
e-mail: {f.ishizawa,tatsuo}@dcl.cs.waseda.ac.jp

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_2

the current most AR only superimposes virtual images or textual information spatially on the real space, typical AR applications are still only human decision-making supports or new types of entertainments. The potential power of AR makes it possible to change the semiotic meaning of the real space and deeply influence human attitudes and behaviors for a better lifestyle [4]. Typical AR approaches can simply integrate computer-generated (CG) virtual images on the real space spatially. The images typically present invisible information that cannot be seen in the actual real world by a user. It is possible to show fictional images to enhance the meaning of the real space, but the approach limits the potential power because only some pieces of objects in the real space can be enhanced. In contrast, virtual reality (VR) technologies make it possible to create a new fictional virtual space by using 3D models of real persons and landscapes; it is possible to present fictional events in the virtual space as people believe that they actually happened in the real space.

This paper proposes a new concept named *Alternative Reality*. In *Alternative Reality*, a user wears a head mounted display (HMD), and the virtual space watched by the user contains fictional events consisting of real landscapes and real persons constructed as a 360-degree movie; thus, the user can navigate the virtual space interactively for offering a new type of ambient intelligence experiences. Because the virtual space is temporally connected and seamlessly integrated with the real space, as shown in Section 3, a user feels that fictional events in the virtual space are performed in the real space even though some events in the virtual space are not real by temporally replacing some real scenes with virtual scenes that contain the fictional events. The approach is a new way to enhance the meaning of the real space to influence human attitude and behavior. Our approach is different from traditional AR techniques that spatially superimpose virtual images on the real space; *Alternative Reality* temporally replaces some real scenes with virtual scenes as presented above. Thus, the approach offers more flexible techniques to integrate virtuality in the real space seamlessly.

In this paper, after presenting an overview of *Alternative Reality*, we demonstrate its basic concept. To show the feasibility of the concept, we developed a case study in which real persons perform some fictional activities in a real location. We conducted a preliminary user study to extract useful insights in order to exploit the feasibility of the *Alternative Reality* concept for developing future ambient intelligence services.

The remainder of the paper is organized as follows. In Section 2, we present several related researches to the proposed approach. Section 3 introduces the *Alternative Reality* concept. In Section 4, a case study of *Alternative Reality* and its experimental results are shown. Section 5 presents future directions of the research, and we conclude the paper.

2 Related Work

Our approach is also similar to the *Substitutional Reality* (SR) system [15]. In the SR system, people's reality is manipulated by allowing them to experience live scenes in which they are physically present and recorded scenes that were recorded and modified in advance without losing a user's reality. Our approach is more general than the SR system because *Alternative Reality* can use CG images generated by VR techniques not only recorded images.

Several previous investigations used a 3D model composed from real scenes. For example, in [2], a user interacts with the 3D model of a building to learn routes inside the building. The user can learn the real routes in the real town in the virtual space. *Alternative Reality* similarly uses the virtual space composed from real persons and landscapes, but the virtual space may contain possible events in the possible future or completely fictional events that may not actually occur in the real space. The events and the stories created from them are perceived as an experience in the real space. By integrating possible future or completely fictional events with real events, a user can accept the events as his/her own real present experience; thus, the presented stories could influence the user's future behavior [12].

AR techniques can be used to enhance existing artifacts. For example, [16] describes several AR games that are enhanced versions of traditional physical games. Specifically, *Augmented Go* [5] demonstrates a promising approach to maintaining the advantages of the physicality of the board game while also adding virtuality. Additionally, *Virtual Aquarium* [8] shows a virtual fish tank, where the movement of the fish reflects a user's toothbrushing behavior. *Enhanced TCG* [14] enhances our real space by replacing a real space component with a fictional component for changing the semiotic meaning of the real space. In these above approaches, a daily object is enhanced by projecting virtual images on it or replacing a real space component with a fictional component. Conversely, *Alternative Reality* temporally replaces some real scenes in the real space with virtual scenes. The approach does not replace a part of a real scene, but replaces the entire images in the real scene with fictional images. It offers a novel possibility to overcome pitfalls described in [14] to integrate the transmedia storytelling approach for incorporating fictionality into the real space [12].

3 Alternative Reality

In *Alternative Reality*, a user watches a sequence of scenes on an HMD. The most important issue in realizing immersion is blurring the boundary between fiction and reality. The use of an HMD offers a better immersive experience by showing a video stream capturing the real space and replacing some real components in the video stream with fictional components [10]. In Fig. 1, there are two persons in the room. One person watches the room through an HMD. Real scene A is a movie that captures the present room. Virtual scene B is a movie that captured the room in advance. The person who wears an HMD watches the real scene at the beginning. Then, the scene are switched to a virtual scene, and he/she can feel that the other person is in front of him/her in the virtual scene. However, in the present real space, the person is actually behind him/her. The sequence that a user watches usually consists of several scenes. Some scenes are captured from present scenes in the real space (Real Scenes in Fig. 1). The scenes are recorded by a camera attached to a user's HMD and shown on the HMD in real-time.[1] However, some scenes in the sequence of the scenes that a user watches are not real-time scenes; the scenes may be actually constructed through VR techniques in advance and are fictional (Virtual Scenes in Fig. 1). Additionally, in the virtual scenes, there are several events that may not occur in the present real space. Typically, the scenes are constructed by 3D models of real persons and landscapes in advance, but some real persons who may not actually present now may appear. One of the important requirements of *Alternative Reality* is that a user feels that these real and virtual scenes are continuous and, thus, is not aware of the boundary between the two scenes. The magic circle is defined as the boundary between the real space and the virtual space [10]. If a user is aware of no magic circle between the two spaces developed by *Alternative Reality*, the user cannot notice that the virtual space generated by *Alternative Reality* is not real. Therefore, he/she feels that the virtual scenes happen actually in the real space.

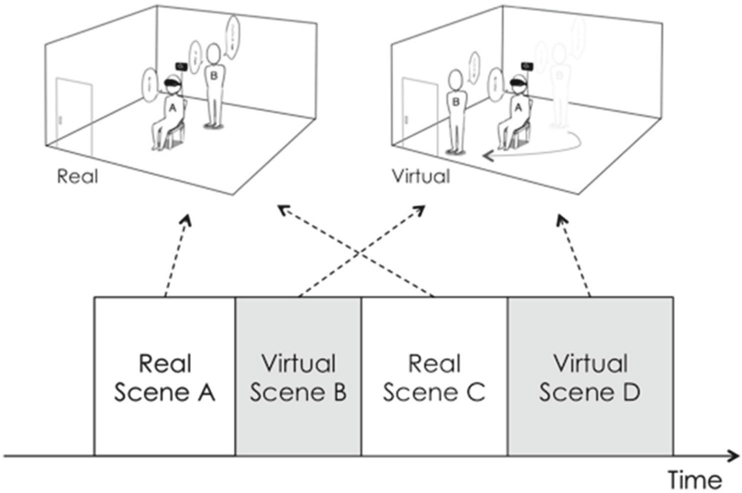


Fig. 1 A Timeline Model of Alternative Reality

[1] In this experiment, a camera is deployed behind a user, and only the center of the captured image is trimmed and shown on his/her HMD. When the user tilts his/her head, the trimmed area shown on the HMD moves according to the movement of the head. This approach simulates a 360 degree movie.

Here, we show some motivating examples of the benefits of adopting the *Alternative Reality* concept:

- Overcoming Social Problems: *Alternative Reality* allows people to be easily aware of some social problems, such as environmental or health issues, because replaced virtual scenes can ironically or exaggeratedly represent the problem. The approach influences people's attitude and behavior. For example, people can explicitly see the impact of his/her unsustainable behavior in the real space by showing damaged natures in the real landscape.
- Serious Game: These games are designed for solving problems in the real space, such as education or mental treatment, rather than entertainment. *Alternative Reality* could be used to develop such games effectively because a player can be more emotionally encouraged in the game. For example, *Alternative Reality* allows a serious game to use real space components for constructing the game world and the game play affects a player's behavior because the player believes that the effect in the game world is realistic due to the use of the real space components.
- Pervasive Game: Extending a user's experience from the pervasive game's space to the real and virtual hybrid space would be easier with *Alternative Reality* [6]. The hybrid space becomes more realistic through *Alternative Reality* because it is easy to insert realistic virtual scenes in the real space seamlessly.

4 A Case Study to Demonstrate Alternative Reality

To demonstrate the *Alternative Reality* concept, we developed a case study. This is still a preliminary stage to demonstrate the proposed concept's feasibility, but the case study allows us to extract useful insights to use the *Alternative Reality* concept for developing more practical services in the future.

4.1 Fictional Future

"*Fictional Future*" has been developed as a case study, that is a movie based on the *Alternative Reality* concept containing both present and possible future occurrences. The case study was developed to extract some sufficiently useful insights and potential pitfalls of the *Alternative Reality* concept. The case study is performed by real known persons and in real locations that participants know to make a user believe that these occurrences are real.

In the "*Fictional Future*" experiment, a participant watched three movies that were merged into one movie through an HMD (Oculus Rift[2]). The first movie is a real-time captured movie that shows the current real location in detail. The second and the third movies are constructed based on the *Alternative Reality* concept, and the scenes in the movies are scenes performed in the past as a possible future.

[2] https://www.oculus.com/

The second movie was performed by an experimenter and captured the scene of the experimenter and participant entering the room. The third movie was performed by the experimenter and consisted of the possible future occurrences to be performed by the experimenter. Fig. 2 shows the screenshots of the second and third movie.

The first photo shows the second movie, where a participant and an experimenter entered a room; this is a past scene. The second photo shows the third movie, where the lady moved ahead; this is a possible future scene. The last photo also shows the third movie, where the lady explained the occurrences; this is also a possible future scene. After watching the merged movie, when the participant took off the HMD, the experimenter performed the same activities performed in the third movie.

Fig. 2 Screenshots of Fictional Future

4.2 Fictional Future Experiment

Eleven participants participated in the "*Fictional Future*" experiment, and we conducted semi-structured interviews with the participants. In the interviews, we asked participants whether they feel that the possible future in "*Fictional Future*" was actually happened in the real space. If their answers are positive, we also asked them the reasons why they felt the reality in the possible future.

Some of them said, "*After taking off the HMD, when I was seeing again the scenes shown in the third movie, I felt déjà vu to see the same scenes again.*" This means that the participants experienced the merged movie made by the *Alternative Reality* concept as real scenes and that after they removed their HMDs, they felt that they had seen the scene before. Additionally, some participants said "*I felt and expected that the activities in the future movie will happen because the activities are desirable and happy for me*". Similarly, after the experiment, a participant asked "*Can I get some money?*" This is because we explained money for gratitude in the third movie; this comment means that the participant believed the possible future presented in the third movie.

An interesting finding of the experiments is that people tend to think that the performed events are real if the activities are common or desirable. That allows us to believe that the events occur in the present real space. Additionally, the realistic landscape is a key to making the events realistic. However, of course, it is important to ensure a sense of the reality of the virtual scenes. If a user loses a sense of the reality, the user confuses what happens in his/her present life and is aware of the boundary of the magic circle because he/she may feel that even real scenes are not real.

A user tends to remember remarkable situations. When a participant feels déjà vu in the experiment, it is typical that these events, such as putting a doll on a desk, touching a doll, or leaving a room, are noticed.

5 Conclusion and Future Direction

In this paper, we proposed a new concept named *Alternative Reality. Alternative Reality* enables us to change our attitude and behavior by presenting some virtual scenes among real scenes. The virtual scenes are constructed from real space components so that a user believes that the virtual scenes actually happen in the real space.

We have developed a case study based on the *Alternative Reality* concept named *Fictional Future* to demonstrate the feasibility of our proposed approach, and we presented some insights extracted from the experiment of the case study.

There are several future directions to enhance the current *Alternative Reality*.

- Realistic Synthetic Space: To exploit the potential of Alternative Reality, the synthetic space needs to be generated by real-time CG. The 3D precise models of performers and environments are also required.

- Social Presence: When we communicate with people in the virtual space, the means of communication should be consistent with communication in the real space [3].
- Feedback: Haptic feedback is important to feel the realistic presence. In addition, if a user wants to walk around the space while using *Alternative Reality*, locomotion is also important [2]. For example, in popular VR systems, a user needs to move using a controller or a keyboard, which does not occur in the real space. One of the problems of using an HMD is motion sickness [7].
- Persuasiveness: If embodied fictionality is natural and realistic, a user may not be aware of the fictionality. For example, when a message is naturally appeared in the user's surrounding, he or she may not notice the message without an explicit feedback. The balance between naturalness and unnaturalness is an important design tool to increase the persuasiveness of the embodied fictionality.

We also need proper augmenting abstractions to enhance the meaning of the real space to influence human attitude and behavior and. Digital rhetoric is a promising direction to offer abstractions to design the enhanced meaning of the real space [13].

References

1. Azuma, R.T.: A survey of augmented reality. Presence **6**(4) (1997)
2. Bailey, J.H., Knerr, B.W., Witmer, B.G.: Virtual Spaces and real world places: transfer of route knowledge. International Journal of Human-Computer Studies **45**(4) (1996)
3. Friedman, F., Widenfeld, H.R, Antley, A., Brongni, A., Slater, M., Garau, M.: Temporal and Spatial Variations in Presence: Qualitative Analysis of Interviews from an Experiment on Breaks in Presence. Presence **17**(3) (2008)
4. Ishizawa, F., Takahashi, M., Irie, K., Sakamoto, M., Nakajima, T.: Analyzing augmented real spaces gamifed through fictionality. In: Proceedings of the 13th International Conference on Advances in Mobile Computing and Multimedia (2015)
5. Iwata, T., Yamabe, T., Nakajima, T.: Augmented reality go: extending traditional game play with interactive self-learning support. In: Proceedings of the 17th IEEE Conference on Embedded and Real-Time Computing Systems and Applications, pp. 105–114 (2011)
6. Jaakko, S., Annika, W., Markus, M.: Pervasive Games. Theory and Design. Experiences on the Boundary Between Life and Play. Morgan Kaufmann (2009)
7. Knerr, B.W., Goldberg, S.L., Bliss, J.P. Moshell, J.J., Blau, B.S. Lampton, D.R.: The virtual environment performance assesment battery: development and evaluation. Technical Report 1029, United States ARMY Research Institute for the Behavioral and Social Science (1995)
8. Nakajima, T., Lehdonvirta, V.: Designing Motivation in Persuasive Ambient Mirrors. Personal and Ubiquitous Computing **17**(1), 107–126 (2013)
9. Nieuwdorp, E.: The pervasive interface: tracing the magic circle. In: Proceedings of DiGRA 2005 Conference: Changing Views – Worlds in Play (2005)

10. Rolland, J., Biocca, F., Hamza-Lup, F., Yanggang, H., Martins, R.: Development of Head-Mounted Projection Displays for Distributed, Collaborative, Augmented Reality Applications. Presence: Teleoperators & Virtual Environments **14**(5) (2005)
11. Sakamoto, M., Nakajima, T.: Incorporating fictionality into the real world with transmedia storytelling. In: Proc. of the International Conference on Design, User Experience and Usability (2015)
12. Sakamoto, M., Nakajima, T.: In search of the right design abstraction for designing persuasive affordance towards a flourished society: In: Proceeding of the 9th International Conference on Design and Semantics of Form and Movement (2015)
13. Takahashi, M., Irie, K., Sakamoto, M., Nakajima, T.: Incorporating fictionality into the real space: a case of enhanced TCG. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (2015)
14. Wakisama, S., Fujii, N., Suzuki, K.: Substitutional Reality System: A Novel Experimental Platform for Experiencing Alternative Reality. Scientific Reports, 2012. doi:10.1038/srep00459 (2012)
15. Yamabe, T., Nakajima, T.: Playful Training with Augmented Reality Games: Case Studies Toward Reality-oriented System Design. Multimedia Tools and Application **62**(1), 259–286 (2013)

Detecting Social Interactions in Working Environments Through Sensing Technologies

Juan Antonio Álvarez-García, Álvaro Arcos García, Stefano Chessa, Luigi Fortunati and Michele Girolami

Abstract The knowledge about social ties among humans is important to optimize several aspects concerning networking in mobile social networks. Generally, ties among people are detected on the base of proximity of people. We discuss here how ties concerning colleagues in an office can be detected by leveraging on a number of sociological markers like co-activity, proximity, speech activity and similarity of locations visited. We present the results from two data gathering campaigns located in Italy and Spain.

Keywords Social network analysis · Computational social science · Smart office

1 Introduction

The analysis of social ties among humans is an important area of research due to its usefulness in understanding the dynamics of human relationships, and, more recently, to its potential in applications like online social networks. In recent years, it has proven its importance also in mobile social networks (MSN), where ties are used to detect communities of users that, in turn, are exploited to improve the strategies for the diffusion of information and services in such networks [1].

J.A. Álvarez-García(✉) · S. Chessa · L. Fortunati · M. Girolami
Department of Computer Science, University of Pisa, Largo Pontecorvo 3, 56127 Pisa, Italy
e-mail: jaalvarez@us.es

S. Chessa · L. Fortunati · M. Girolami
Italian National Council of Research (CNR-ISTI), Via G. Moruzzi, 1, 56124 Pisa, Italy
e-mail: {stefano.chessa,luigi.fortunati,michele.girolami}@isti.cnr.it

J.A. Álvarez-García · Á.A. García
Computer Languages and Systems Department, University of Seville, 41012 Seville, Spain
e-mail: aarcos1@us.es

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_3

Ties can be detected on the base of contextual information, either from users' agendas, telephone calls registries or e-mails [2], or from sensors [3]. Recent approaches in MSN identify the ties on the base of co-location traces among users [3, 4], mainly for two reasons:(i) co-location in MSN is associated to frequent and lengthy contacts among users, which are essential to sustain the exchange and diffusion of information; and (ii) co-location is easily producible by using short-range radio interfaces (such as Bluetooth and WiFi in ad-hoc mode) that are widely available on smartphones and other personal devices. MSN use ties to detect communities, and assume that users in the same community share the same interests. This assumption is very useful to reduce the overhead of the information diffusion strategies since it avoids passing information to users that are not interested in it, and it maximizes the probability of passing the information to users that will most likely meet with other interested users. However, the progress of research and experimentation in this field showed that this assumption is usually unmet (for example, two users that take the same metro every day may share no interests at all). Following this observation, a number of recent works proposed to detect ties based on a richer set of contextual information [4, 5, 6, 7].

This work focuses on detection of ties in MSN and it fits this last trend of research. Specifically, we aim at finding ties based on a short-term analysis of information about the user's actions and interactions detected by the sensors embedded in conventional smartphones. Short-term analysis means that ties are detected on the base of short temporal sequences (a few minutes at most), in order to provide timely information to the community detection algorithms used to support the information diffusion in MSN. Specifically, we limit this preliminary study to ties among colleagues from the same office. The novel aspects of our approach stands in the identification of a number of activities beyond co-location (also referred to as proximity) that are related to a deeper interaction between two users (and thus to a stronger tie between them). Namely, we focus on symmetry of the physical activity of users, i.e. activities that two users are carrying out together, like walking side-by-side or resting and we relate this to the semantic of the place where activities are being carried out (for example the office, a relaxing area, the cafeteria or a meeting room). To this purpose, we set up an experiment in two sites (at CNR located in Pisa and at the Department of Computer Science in Seville) and that involves 8 participants for each location. Participants are equipped with a smartphone running a sensor data logger app. Moreover, they also keep track of their interactions with other colleagues during the working day. This information is used to assess the ground truth of the interactions against which we compare our algorithm for detecting ties. This paper presents preliminary results on the analysis of a selection of sociological markers that we considered meaningful for detecting the existence and the intensity of the social interactions among colleagues.

2 Related Works

The recent advances in the analysis of complex social networks have recently stimulated research efforts from a rather wide range of viewpoints in social studies through sensing devices [8, 9]. There is a large body of proposals related to the integration of social networks and pervasive computing, whose main objective is to analyze social interactions and organizational structures among individuals, starting from a set of sensing solutions.

The availability of a set of sensors on mobile devices, such as cameras, GPS, microphones and accelerometers, along with short-range communication interfaces such as Bluetooth and WiFi [4], enables new sensor-based applications oriented towards activity recognition and interactions with other users. In [3], for example, the authors propose a probabilistic model to discover several interaction types from large-scale dyadic data such as proximity of personal smartphones (via Bluetooth co-location traces), phone calls and emails. Using Bluetooth as a proximity sensor to reconstruct social dynamics at large scale has also been investigated in the reality mining initiative [4]. This project studies complex social systems by inferring patterns in daily user activity, relationship among users, meaningful social locations, and organizational structures. The reality mining dataset was also used to identify social groups and to infer frequency or duration of meetings within each group. Another work [10] showed that it is possible to infer different interaction types using a probabilistic model applied on Bluetooth data.

Following studies on proxemics [11], researchers have also employed mobile phones [12] or custom designed personal devices, such as sociometric badges [13], to capture in detail users body movements, relative body orientation, interpersonal distance or recurrent speech patterns, as a mean to detect face-to-face interactions. The study of social interactions in academic or work environments proved to be an interesting topic in the research community. These works are mainly motivated by the beneficial effect that social interactions have in collaborative environments, driving the innovation process [14] or generally improving the well-being of employees [15].

3 Modelling Social Interactions in Mobile Social Networks

The network of mobile devices carried by people is a MSN that can be modeled as a temporal graph $G_t = (N, E_t)$ where N is the set of devices in the network and E_t are the ties among devices at time *t*. An edge $e_{i,j} \in E_t$ is the tuple $e_{i,j} = (n_i, n_j, w)$, where w quantifies the intensity of the tie between devices n_i and n_j. We define $e_{i,j}$ active if and only if $w > 0$. Graph G_t is the snapshot of the MSN at time *t*. Given an observation interval starting at time t_0 and ending at time t_k the sequence of snapshots $G_{[t_0,t_k]} = \{G_{t_0} \cdots G_{t_k}\}$ is the time-varying graph describing the evolution of the ties among people. We compute the weight w among devices with the function γ that evaluates simultaneously a number of *sociological* *markers* extracted during an observation interval. Some examples of markers are the locations visited by people, the voice activity or the physical activity of people. The combination of them forms the weight of the tie of a dyad, as described with the following general definition:

$$w = \gamma_{[t_0,t_k]}(n_i, n_j) \quad (1)$$

We observe that two individuals are involved in a social relationship if some sociological markers appear evident during a period. For example, shaking hands, talking to each other, staying close during a meeting underlie the explicit willingness of interaction between them. Such markers are evaluated by analyzing the output of sensors available on the people's devices with a given sampling rate λ. Hereafter, we consider the following sociological markers and we assume that they are all analyzed at the same rate.

The proximity marker reveals which devices lay within the *intimate* zone of a person, (for example, the range can be of 10 meters like in Bluetooth). If device n_i detects n_j ($n_i \rightarrow n_j$) then we can assume that the device's owners are also close and they might being interacting. We model the proximity with the Boolean function $f_P(n_i, n_j)$ which is 1 if n_i *or* n_j are in range of each other, and 0 otherwise.

The speech marker is used to detect if people are talking. We are not interested in recognizing the voice of a person or in recognizing the meaning of the speech, rather we limit to detect if people are talking. We model the speech as the function $f_S(n_i, n_j)$ which is 1 if either n_i *or* n_j are talking, and 0 otherwise.

We recognize the physical activity of two persons with the goal of determining if there exists symmetry, in particular if they are doing the same activity. When combined with other markers, the symmetry of activities is an amplifier of the intimacy of the relationship between a pair. For example, detecting that two people are in proximity, and that they are talking and walking is a stronger combination than detecting only proximity. Among the physical activities (walking, cycling, running or tilting) we only consider two categories: moving or still. We model the activity with the function $f_A(n_i, n_j)$ which is 1 is n_i *or* n_j are both still or moving, and 0 otherwise

The last marker measures the similarity of the locations visited by a pair of individuals. The similarity of the visited locations is given by the similarity of the WiFi access points detected along the time. The more the list of the access points of a pair is similar, the more likely they visited the same locations and hence the stronger is the intimacy of the relationship between them. Given the list L_i of access points stored by a device, we compute the similarity of locations with the Adamic-Adar [16] score: $f_L(n_i, n_j) = \sum_{l \in L_i \cap L_i} \frac{1}{\log |hit_i(l) + hit_j(l)|}$. The Adamic-Adar metric considers only the locations l detected from both devices, and it returns the inverse of the logarithmic of the number of times n_i and n_j detected l. The inverse of the logarithmic penalizes very popular locations that might be not so representative of an intimate interaction.

4 Experimental Settings and Results

We analyze the social interactions among people with a number of participants from Italy and Spain. *SmartRelationship* is the Android data logger application designed to access the device's sensors and store the collected data. All the data is collected and analyzed at the end of the experiment to detect social ties between different users. As described in Section 3, we consider a number of markers, namely proximity, speech, physical activity and location that are all sampled with a rate $\lambda = 30s$. During the working hours (9 A.M. to 6 P.M.), *SmartRelationship* stores the following information:

Proximity: the set of Bluetooth MAC addresses, the name and the category of the devices lying within the range of 0 to 10 meters are stored. The application performs one Bluetooth scan every λ seconds. We then post-process the proximity logs by excluding the MAC addresses of devices not recognized as part of the experiment.

Speech: five-second audio clips with the device's microphone are recorded. The goal is to detect voice or silence, disregarding the meaning and source of the speech. The audio clips are post-processed through a voice activity detection algorithm based on Gaussian Mixed Model, which detects if people are talking or not. This algorithm does not discriminate between noise and voices so it is not a key feature in our system. The solution is naturally prone to false positives, since speech can be detected from external conversations not involving the user.

Activity: the physical activity of the device's owner by exploiting the Google Activity Recognition APIs is gathered. The list of detected activities is filtered by three means of activities: in vehicle, on bicycle and on foot. If the sum of their confidence is over 50%, then the activity detected is *moving*, otherwise it is *still*.

Location: information about the WiFi access points detected is gathered. In particular, we store the WiFi network name (SSID), the MAC address of the access point (BSSID) as well as the RSSI of the access point.

We conducted the experimentation in two locations, namely the ISTI-CNR in Pisa (Italy), and the University of Seville (Spain). In each location we recruited 8 volunteers. The experiment lasted for 5 days, from October 26th to October 30th, 2015. During each day, each participant filled a questionnaire to keep track of all his meaningful social interactions, annotating the people involved and the time window with one-hour granularity. To our purposes, an interaction is meaningful if it lasts at least 2 minutes and it is a face-to-face interaction involving at least proximity and voice. This excludes short and involuntary interactions with other colleagues.

We analyze the dataset obtained by post-processing the logs produced by *SmartRelationship*. Note that we downsample the dataset at $\lambda' = 120s$ because the devices of the participants shown very different performances, and slower devices could not meet a sampling rate of 30 seconds. Also, social interactions evolve following the human rhythms and we consider that $\lambda' > \lambda$ is a reasonable observation period.

For each time interval of width λ' seconds, we aggregate all the samples within such interval. Figure 1 shows how the markers are grouped. The blue dotted line shows the proximity marker (f_P) with rate λ, the orange line shows f_P with $\lambda' = 120s$.

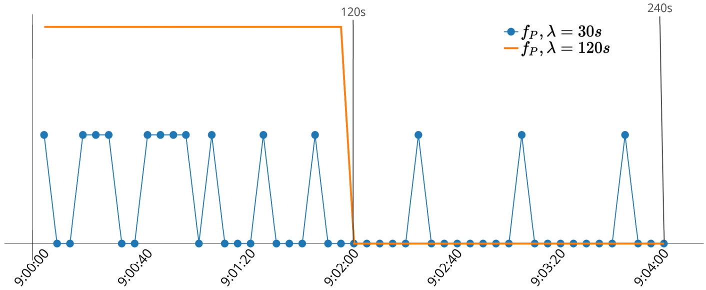


Fig. 1 Sampling rate of the proximity marker.

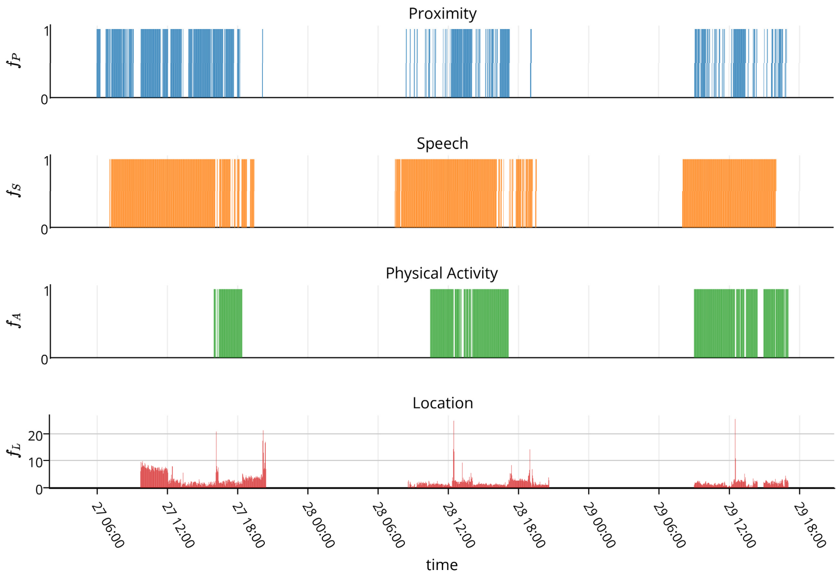


Fig. 2 Time series of the sociological markers of two participants.

We first analyze the time series of the sociological markers described in Section 3 to identify those features underling a social interaction. Figure 2 shows the time series of the proximity, speech, activity and location markers of two participants for 3 days.

In the figure, the proximity marker follows an intuitive but important pattern. In fact, contacts among participants happen during the daily working hours (from 9.00 AM to 6 PM). On the other hand, when $f_P = 0$ then participants are not in proximity and hence they are not involved in any social interaction, differently when $f_P = 1$ it means that they are close and they might interact. The speech marker is useful to identify the time periods of absence of verbal interactions (when $f_S = 0$). In fact, if two people are in silence we assume that they are not interacting. The activity and the location markers are used as amplifier of the magnitude of the interactions. In particular, when we do not detect symmetry of physical activity (i.e. when $f_A = 0$), then we penalize the intensity of the tie, conversely when $f_A = 1$ we increase the intensity. Lastly, the location marker reveals the similitude of locations visited along the time. An interesting property from the location marker is that colleagues spending most of the time in their office have a low score of the location marker, because the locations detected become more and more popular day after day. Differently, as soon as two participants move toward a new location, the location score increases.

We then combine together the sociological markers shown in Figure 2 with the goal of assigning a weight w to the interactions among participants along the time. To this propose we define the γ function (defined in a general form in (1)) as follows:

$$\gamma_{[t_0,t_k]}(n_i, n_j) = f_P \cdot [f_S \cdot (f_A + \log(f_L))]$$

Figure 3 shows the results of the γ function computed for the time series shown in Figure 2. We compare the γ function against the ground truth obtained by analyzing the questionnaires of the participants. Since questionnaires are not symmetric (the answers from the participants might not be the same), we assign a higher value to the time slots in which we found concordance of answers and lower values when we detect discrepancies on the answers.

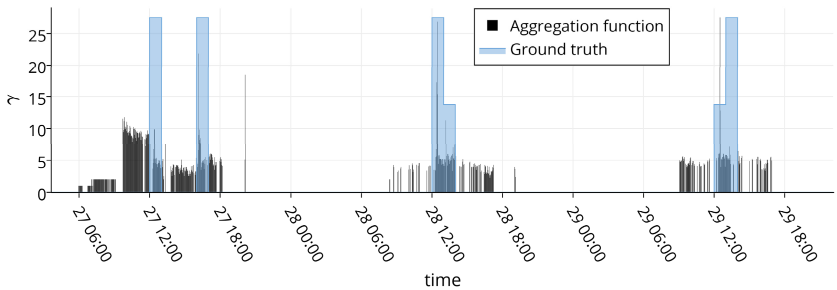


Fig. 3 The fusion algorithm applied to the sociological markers of two participants.

Figure 3 shows that the intensity of the interactions between the two participants changes along the time. When the proximity and/or speech markers are not active (f_P and f_S are 0), then we do not detect interaction and we assign $\gamma = 0$. This is the case for Tuesday 27th from 10.30 to 12.00 and sparsely on Wednesday 28th before 12.00). Differently, $\gamma > 0$. In particular, γ increases when the two participants are in proximity, they are talking and they are visiting a new location (during lunch time at the canteen).

5 Conclusions

The identification of social ties and the assessment of their strength is becoming a problem of great interest due to its application to MSN. Recent works have already identified the limit of considering only co-location to identify ties. This work follows this trend of research and, on the base of a preliminary experimentation on a case of colleagues in an office, we analyze the information provided by markers concerning co-activity such as talking or moving along with location and the more conventional proximity. Although only two subjects are studied, to check the validity of our implementation, we can conclude that similarity in physical activity and fusion of several sources help us to detect real social interactions. Future work will expand this experimentation and will test the method for identification of communities.

Acknowledgments This research is partially supported by the projects of the Spanish Ministry of Economy and Competitiveness HERMES (TIN2013-46801-C4-1-r) and Simon (TIC-8052) of the Andalusian Regional Ministry of Economy, Innovation and Science.

References

1. Girolami, M., Chessa, S., Caruso, A.: On Service Discovery in Mobile Social Networks: Survey and Perspectives. Computer Networks **88**, 51–71 (2015)
2. Vanini, S., et al.: Modeling social interactions in real work environments. In: Proc. 6th International Workshop on Hot Topics in Planet-Scale Measurement. ACM (2015)
3. Do, T.M.T., Gatica-Perez, D.: Human interaction discovery in smartphone proximity networks. Personal and Ubiquitous Computing **17**(3), 413–431 (2013)
4. Eagle, N., Pentland, A.: Reality mining: sensing complex social systems. Personal and Ubiquitous Computing **10**(4), 255–268 (2006)
5. Matic, A., Osmani, V., Mayora-Ibarra, O.: Analysis of social interactions through mobile phones. Mobile Networks and Applications **17**(6), 808–819 (2012)
6. Junker, H., et al.: Gesture spotting with body-worn inertial sensors to detect user activities. Pattern Recognition **41**(6), 2010–2024 (2008)
7. Wyatt, D., et al.: Inferring colocation and conversation networks from privacy-sensitive audio with implications for computational social science. ACM Trans. on Int. Systems and Technology (TIST) **2**(1), 7 (2011)

8. Lane, N.D., et al.: A survey of mobile phone sensing. Communications Magazine, IEEE **48**(9), 140–150 (2010)
9. Stopczynski, A., Sapiezynski, P., Pentland, A., Lehmann, S.: Temporal fidelity in dynamic social networks. Eur. Phys. J. B **88**, 249 (2015)
10. Do, T.M.T., Gatica-Perez, D: Groupus: smartphone proximity data and human interaction type mining. In: 15th Sym. on Wearable Computers (ISWC) (2011)
11. Hall, E.T.: Proxemics: The study of man’s spatial relations (1962)
12. Banerjee, N., et al.: Virtual compass: relative positioning to sense mobile social interactions. In: Pervasive Computing, pp. 1–21. Springer, Heidelberg (2010)
13. Onnela, J.-P., Waber, B.N., et al.: Using sociometers to quantify social interaction patterns, Scientific Report 4 (2014)
14. Peschl, M.F., Fundneider, T.: Designing and enabling spaces for collaborative knowledge creation and innovation: From managing to enabling innovation as socio-epistemological technology. Computers in Human Behavior **37**, 346–359 (2014)
15. Dimotakis, N., Scott, B.A., Koopman, J.: An experience sampling investigation of workplace interactions, affective states, and employee well-being. Journal of Organizational Behavior **32**(4), 572–588 (2011)
16. Adamic, L., Adar, E.: Friends and neighbors on the web. Social Networks **25**(3), 211–230 (2003)

Estimation of the Optimum Speed to Minimize the Driver Stress Based on the Previous Behavior

Victor Corcoba Magaña, Mario Muñoz Organero, Juan Antonio Álvarez-García and Jorge Yago Fernández Rodríguez

Abstract Stress is one of the most important factors in car accidents. When the driver is in this mental state, their skills and abilities are reduced. In this paper, we propose an algorithm to predict stress level on a road. Prediction model is based on deep learning. The stress level estimation considers the previous driver's driving behavior before reaching the road section, the road state (weather and traffic), and the previous driving made by the driver. We employ this algorithm to build a speed assistant. The solution provides an optimum average speed for each road stage that minimizes the stress. Validation experiment has been conducted using five different datasets with 100 samples. The proposal is able to predict the stress level given the average speed by 84.20% on average. The system reduces the heart rate (15.22%) and the aggressiveness of driving. The proposed solution is implemented on Android mobile devices and uses a heart rate chest strap.

Keywords Intelligent transport system · Stress driver · Driving assistant · Deep learning · Particle Swarm Optimization · Android · Mobile computing

1 Introduction

According to many researchers, excessive speed is the cause of a large number of accidents. Therefore, it is one of the most important variables for predicting the risk of accidents and its severity. In [1], the authors of the report pointed out that

V.C. Magaña(✉) · M.M. Organero
Telematic Engineering Department, Universidad Carlos III de Madrid, 28911 Leganés, Spain
e-mail: {vcorcoba,munozm}@it.uc3m.es

J.A. Álvarez-García · J.Y.F. Rodríguez
Computer Languages and Systems Department, University of Seville, 41012 Seville, Spain
e-mail: {jaalvarez,jorgeyago}@us.es

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_4

speeding (speed higher than speed limits) is responsible for 18% of accidents in London during the year 2014, whereas travelling too fast for the road conditions caused 8 per cent of fatal accidents. Reducing average speed, results in a decrease of the number of fatalities [2].

Besides the vehicle speed, there are other related factors that affect safety such as sudden accelerations and decelerations and aggressive driving. 56% of fatal accidents between 2003 and 2007 involved one or more aggressive actions [3], such as: suddenly changing speeds, making an improper turn or erratic lane changing. This behavior is often related to the stress of the driver. Stress can be defined as a change from a calm state to an excitation state in order to preserve the integrity of the person. If the stress is negative, it is called ‘‘distress’’. This type of stress is commonly due to an increase in the workload, such as traffic density, inappropriate vehicle speed, etc. It causes irritability, lower concentration and problem to take decisions and has also physical symptoms such as headache and a fast heartbeat. Stressed drivers are more likely to engage in risky behaviors and having accidents [4].

There are many proposals to detect stress and measure the workload [5]. Most of them are based on physiological features such as electromyogram, electrocardiogram, respiration, and skin conductance. Non-invasive and non-intrusive sensors are particularly interesting for measuring driver stress.

In the literature there are many assistants that help to adapt the speed in order to maximize safety and reduce fuel consumption. These systems are called “Intelligent Speed Adaptation” (ISA). A large number of researchers have evaluated the benefits that these systems provide regarding safety and fuel consumption. ISA has been evaluated estimating a reduction of the accidents between 25% and 30% if all vehicles use ISA system.

2 Our Approach

In this work, we propose a driving assistant that estimates the optimal average speed for the next road section taking into account the driving and the stress level on previous road section. In addition, the algorithm considers the vehicle telemetry obtained by the driver in some of the previous driving under similar driving conditions.

The idea is that the driving assistant proposes an average speed for the road section, which minimizes the driver stress and does not increase the trip time significantly. This speed does not require a sudden change in driving habits because the proposed average speed is based on the driver. The objective is to avoid one of the biggest problems of the driving assistant: frustration and impatience.

2.1 Estimation of the Optimum Speed to Avoid Stress

Particle Swarm Optimization (PSO) [7] and Deep Learning [8] are used to estimate the average speed for each section of the road. Figure 1 shows the flowchart to estimate the optimum speed. The main advantage of PSO algorithm is that it maintains multiple potential solutions at one time. We employed Deep Learning as fitness function because it allows us to represent complex relationships between data. Other algorithms were tested during the test stage as neural networks. However, deep learning obtained the highest hit ratio.

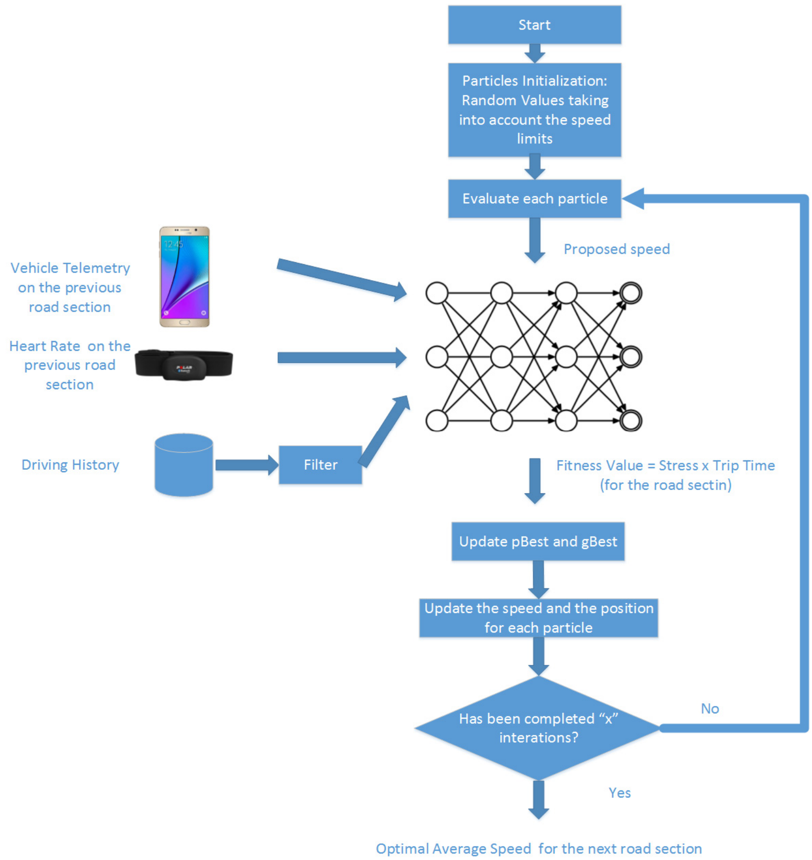


Fig. 1 Flowchart to estimate the optimum speed for reducing stress and fuel consumption

Each solution is represented by a particle in the search space. It has the following elements:

- Position: In our case is the recommended speed for reducing the stress level.
- Pbest: This is the best position on the current particle (speed that minimizes the heart rate and improves the driving style).

- Gbest: It is the best position among all particles.
- Speed: Is calculated using equation 1. It determines which will be the next speed of the particle.

$$v_i(t) = w*v_i(t-1) + c1*r1*(Pbest(t-1) - x_i(t-1)) + c2*r2*(Gbest-x_i(t-1))$$

where $v_i(t)$ is the particle's velocity at time t, w is the inertia weight, $x_i(t)$ is the particle's position at time t, Pbest is the particle's individual best solution as of time t, and gBest(t) is the swarm's best solution as of time t, c1 and c2 are two positive constants, and r1 and r2 are random values in the range [0–1]

The particles "fly" or "swarm" evolve through the search space to find the minimum value. During each iteration of the algorithm, they are evaluated by an objective function to determine its fitness. Next position is calculated by equation 2:

$$x_i(t+1) = x_i(t) + v_i(t+1)$$

where $x_i(t)$ is the current particle's position and $v_i(t + 1)$ is the new velocity.

2.2 Evaluation of the Particles

The fitness function is a deep learning algorithm. This type of algorithms have their origin in neural networks, the main difference is the number of hidden layers and the training process. Deep learning algorithm estimates the stress level given a recommended speed (the particle position) based on:

- The driving and the stress level on the previous road section
- Telemetry samples from the next road section takes by the user other days

In figure 1, we can see that there is a filter applied to the driving history. The algorithm is trained with samples of driving that were obtained under conditions similar to the present ones. We must bear in mind that speed is a variable that depends on a large number of factors such as the traffic state or weekday. In our case, we consider the following elements to filter:

- Weather State: The number of vehicles on the road grows when the weather conditions are bad, increasing the likelihood of traffic incidents. Moreover, the roll coefficient changes. Therefore, the advice have to take into account that factor. In addition, many studies highlight that when it is hot, the fatigue appears before. On the other hand, cognitive capacity of the driver worsens when it is cold.
- Traffic State: When traffic is heavy, the stress level increases. In these cases, the vehicle speed must be adjusted in order to avoid accelerations and decelerations.
- Time: Fuel consumption is increased at rush hours. The driver has to accelerate and slow down more frequently. In addition, the engine is switched on during more time. This situation causes stress, increasing the

accidents risk. On the other hand, night driving maximizes the likelihoods of sleep despite he or she has previously slept. This is due to the sleep cycle. Therefore, we have to take into account the time when we estimate the optimum speed.

- Weekday: The rush hour depends on the day. For instance, it is common that the average speed is higher on weekend because there is less traffic. If the road conditions are good, the system should recommend an average speed high enough to avoid the frustration of the driver.

Deep Belief Network that we have defined in our proposal has the following input variables:

- Recommended average speed (km/h): It is the vehicle speed proposed by the PSO algorithm. If the speed is too high, the stress will increase because driver has less time to make decisions. On the contrary, if the speed is abnormally low, fuel consumption will increase because the engine will be running more time. In addition, it may cause traffic incidents and the stress level from road users would be higher.
- Positive Kinetic Energy (m/s^2): It measures the aggressiveness of driving and depends on the frequency and intensity of positive accelerations [9]. A low value means that the driver is not stressed and drives smoothly. The value is calculated using the following equation:

$$PKE = \frac{\sum(v_i - v_{i-1})^2}{d};\ v_{i-1} < v_i$$

where v_i is the vehicle speed (m/s) and d is the trip distance (meters) between v_i and v_{i-1}.
- Average Acceleration (m/s^2): The acceleration (positive and negative) may indicate the presence of stress or fatigue. These actions increase the likelihood of traffic accidents and fuel consumption.
- Speed (Km/h): This variable allows us to model the driver behavior. For each section, we consider the following measures statistics: minimum, medium, maximum, average, and standard deviation.
- Average Heart Rate (b.p.m): Heart Rate signals are employed as an indicator of ANS neuropathy for normal, fatigued and drowsy states because the ANS is influenced by the sympathetic nervous system and parasympathetic nervous system. This indicator is not intrusive. A high heart rate means the driver has stress.

The output of the deep learning algorithm will be an estimate of the stress level when the driver drives at “x” km/h on average. The output can take the following values:

Low Stress	Normal Stress	High Stress
0.5	1	1.5

The final value of the fitness function is given by the value of estimated stress and the time required to complete the road section. The goal is to find a velocity that balances stress level and travel time, staying safe on the road.

2.3 Experimental Design

The speed assistant was deployed on Galaxy Note 4. This devices supports Bluetooth Low Energy. Heart Rate was monitored using GEONAUTE[1]. In order to validate the algorithm, we employed five datasets with 100 driving samples. Table 1 describes the dataset features. Each column captures the number of samples where the driver stress is low, normal and high, respectively. The driving samples were obtained on real tests. Each dataset contains the vehicle telemetry and the heart rate obtained in the current road section and the previous road section. Driving tests were performed in Seville. Trip distance was 23 km. This track has highway, secondary road, and urban road. The route was divided into section of 500 meters. We have chosen this length after making several tests with the datasets. Figure 2 captures the entire route (green line). We chose five road sections to analyze the algorithm. The vehicle was a Seat Ibiza Sport 1.9 TDI (2008). Tests were performed under similar road conditions.

Table 1 Dataset Features employed to validate the algorithm

	Low Stress	Normal Stress	High Stress	Total
Dataset 1	58	30	12	100
Dataset 2	18	53	29	100
Dataset 3	36	28	36	100
Dataset 4	34	21	45	100
Dataset 5	35	43	22	100

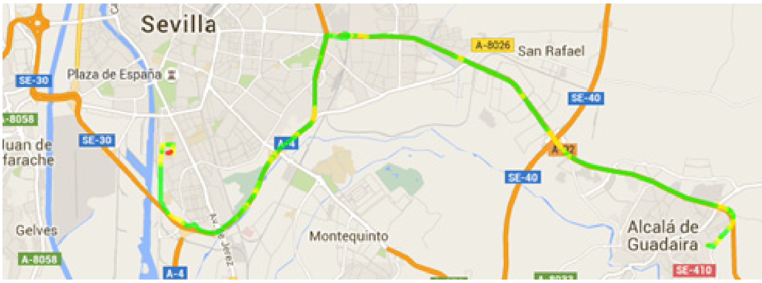


Fig. 2 Route for the driving tests

[1] http://www.geonaute.es/cinturon-cardiofrecuencimetro-bluetooth-smart-40-id_8288269

2.4 Results

Table 2 presents the results (Mean squared error and hit rate) of the algorithm considering the five datasets and using 10 fold cross validation. The proposal is able to predict the stress given the average speed by 84.20 % on average. The deep learning algorithm was run with 3 hidden layers, 200 units per layer and 300 epochs.

Table 2 Results of the 10 fold cross validation using the proposal

	MSE	Hit Rate (%)
Dataset 1	0.156802	84 %
Dataset 2	0.131412	85 %
Dataset 3	0.111304	89 %
Dataset 4	0.071342	92 %
Dataset 5	0.283674	71 %

Figure 3 shows the average heart rate obtained by the driver on different road sections without and with the speed assistant. The heart rate decreases by 15.22 % on average taking into account all driving samples. The heart rate value is influenced by the sympathetic nervous system and parasympathetic nervous system. A low value means that the driver is not stressed.

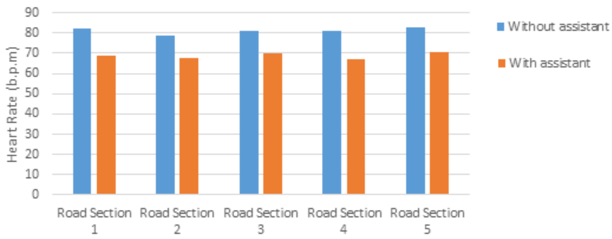


Fig. 3 Average heart rate (b.p.m) using the speed assistant

Figure 4 illustrates the positive kinetic energy obtained by the driver when driving with and without the speed assistant. This variable measures the aggressiveness of driving and depends on the frequency and intensity of positive acceleration. A low value means that the driving is not aggressive. The positive kinetic energy decreases by 33.53 % on average when the driver uses the proposal. Moreover, a smoother driving saves more fuel.

Fig. 4 Positive Kinetic Energy (m/s^2) using the speed assistant

Figure 5 shows the average standard deviation of vehicle speed for each road section. Following the recommended speed provided by the solution, standard deviation is reduced 37.09 % on average.

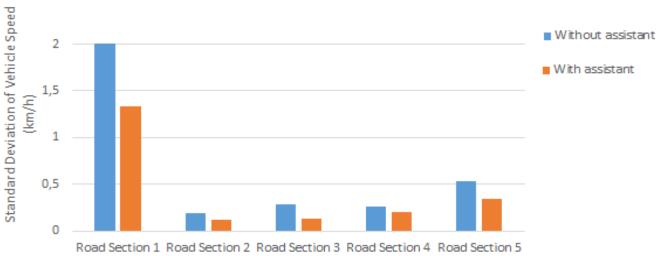


Fig. 5 Standard Deviation of Vehicle Speed (km/h) using the speed assistant

3 Conclusions

In this paper, we have proposed an algorithm to estimate the driver stress given speed and heart rate for a road section. The goal is to prevent stress and dangerous situations, based on current driving context, taking data from road conditions (weather and traffic), time, and previous driving. In this paper, we present a speed assistant that uses driving context, deep learning technics along with Particle Swarm Optimization (PSO) to calculate a speed that minimizes the stress without increasing the trip time. The results show a significant improvement in the driving style and the driver stress.

As future work, we want to validate the algorithm with more users and more routes. It would also be interesting to analyze the effect that the solution has in the stress of the other road users. In addition, we want to introduce new variables to improve the prediction model such as previous stress, working time, sleeping time, and quality of sleep.

Acknowledgments The research leading to these results has received funding from the "HERMES-SMART DRIVER/CITIZEN" projects TIN2013-46801-C4-2-R /1-R funded by the Spanish MINECO, from the grant PRX15/00036 from the Ministerio de Educación Cultura y Deporte.

References

1. Department for Transport. Reported road casualties in Great Britain: main results (2014). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/438040/reported-road-casualties-in-great-britain-main-results-2014-release.pdf
2. Aarts, L., Van Schagen, I.: Driving speed and the risk of road crashes: A review. Accident Analysis & Prevention **38**(2), 215–224 (2006)
3. Aggressive driving: Research update. April 2009, October 2015. http://www.aaafoundation.org/pdf/AggressiveDrivingResearchUpdate2009.pdf
4. Hill, J.D., Boyle, L.N.: Driver stress as influenced by driving maneuvers and roadway conditions. Transportation Research Part F: Traffic Psychology and Behaviour **10**(3), 177–186 (2007)
5. Solovey, E.T., et al.: Classifying driver workload using physiological and driving performance data: two field studies. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM (2014)
6. Lai, F., Carsten, O., Tate, F.: How much benefit does Intelligent Speed Adaptation deliver: An analysis of its potential contribution to safety and environment. Accident Analysis & Prevention **48**, 63–72 (2012)
7. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, vol. 4, pp. 1942–1948, November/December 1995. doi:10.1109/ICNN.1995.488968
8. Sun, N., Han, G., Du, K., Liu, J., Li, X.: Person/vehicle classification based on deep belief networks. In: 2014 10th International Conference on Natural Computation (ICNC), pp. 113–117, Augest 19–21, 2014. doi:10.1109/ICNC.2014.6975819
9. González, R., Belén, A., et al.: Modeling and detecting aggressiveness from driving signals. IEEE Transactions on Intelligent Transportation Systems **15**(4), 1419–1428 (2014)

Using Topic Modelling Algorithms for Hierarchical Activity Discovery

Eoin Rogers, John D. Kelleher and Robert J. Ross

Abstract Activity discovery is the unsupervised process of discovering patterns in data produced from sensor networks that are monitoring the behaviour of human subjects. Improvements in activity discovery may simplify the training of activity recognition models by enabling the automated annotation of datasets and also the construction of systems that can detect and highlight deviations from normal behaviour. With this in mind, we propose an approach to activity discovery based on topic modelling techniques, and evaluate it on a dataset that mimics complex, interleaved sensor data in the real world. We also propose a means for discovering hierarchies of aggregated activities and discuss a mechanism for visualising the behaviour of such algorithms graphically.

1 Introduction

The goal of activity recognition (AR) is to automatically identify the activities of humans using data collected from networks of sensors. The sensors may be embedded in the environment or can be wearable sensors. The dominant approach to AR is to train activity identification models by running machine learning algorithms on annotated datasets [10]. Typically these datasets have to be hand annotated. This can be problematic in terms of the human labour required to create a dataset and in the accuracy of the annotations.

Activity discovery (AD) has been proposed as an alternative to manual annotation. AD involves the unsupervised discovery of activities from raw sensor data. In this paper we present an approach to AD based on topic modelling. The approach can discover activities and also discover activity hierarchies where high-level complex

E. Rogers(✉) · J.D. Kelleher · R.J. Ross
Applied Intelligence Research Centre, Dublin Institute of Technology, Dublin, Ireland
e-mail: eoin.rogers@student.dit.ie, {john.d.kelleher,robert.ross}@dit.ie

© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_5

activities are composed of simple activities. The approach is implemented in a system that is tested on a dataset that mimics complex, interleaved sensor data. As a second form of evaluation we propose a graphical visualisation of the output of the system which is designed to provide insight into the hierarchies of activities that the system identifies.

Paper structure: In Sect. 2 we review prior work in the area of activity discovery. Then, Sect. 3 describes our approach to AD. Finally, Sect. 4 presents an evaluation of our system and makes some concluding remarks.

2 Prior Work on Activity Discovery

AD is concerned with the identification of activities in a dataset independently of human annotation [3]. An AD system generally works by identifying recurring patterns in the sensor data without attempting to understand what activity the pattern signifies (this semantic information has to be added by a human user at a later date). AD is already a well-established field of research and as a result a number of approaches to the task have been proposed and tested.

Cook et al. [3] notes that many pattern mining algorithms have been used without modification, and with some success, for AD. [3] also propose their own algorithm which uses beam search to discover patterns that best compress the data from candidate patterns generated by extending previously discovered patterns. Genetic algorithms are an interesting way of tackling the activity discovery problem. The genetic algorithm proposed in [11], called GAIS, uses comparisons to pre-discovered patterns (called "*models*" by the authors) as a fitness function. This makes this approach well-suited for detected variations of existing patterns, but it is not optimised for detecting novel patterns. Likewise, Nguyen et al.'s SuperAD system [8] trains machine learning models on a partially labelled dataset, and then takes events that cannot be recognised with high accuracy by the system to be discovered events, which are presented to the user for manual labelling. Some authors have proposed more novel means of discovering activities, for example the system proposed by [7]. One interesting aspect of this system is the use of the *Minimum Description Length* (MDL) principle to evaluate performance.

Another approach in the literature to AD is topic modelling, for example as proposed by [6], in which a sliding window over temporal sensor data is used to produce "*documents*" that can be fed into a topic modelling algorithm. For example, given a dataset consisting of sensor readings over a one hour period, the user selects a temporal period of, say, 120 seconds and an increment of, say, 20 seconds. A sliding window is placed over the dataset starting at 0 seconds and ending at 120 seconds. The contents of the window are copied into a document, and the window is incremented by 20 seconds (so it is now covering the period from 20 seconds to 140 seconds). Another document is produced by copying the window's contents, and again the window is incremented. This process continues until the end of the dataset is reached. This will produce a number of documents (actually $n-w+1$ documents, where n is the length of the dataset and w is the window length).

The algorithm proposed by [6] is conceptually simple and it is able to detect novel patterns in previously unseen data. Nonetheless, we take the view that there are a number of issues with the approach taken by this paper that could be improved upon. Firstly, as noted by [5], the dataset used did not contain any *interleaving* (in other words, there were no situations where multiple activities occurred at the same time). Since most real-word AD systems will be expected to deal with interleaved data, we feel that this substantially limits the usefulness of the results presented. On top of that, the sensors used in the original papers were wearable sensors attached to the subject's body. Such sensors can collect detailed information about the user, since they can potentially record every action they carry out. However, many potential users are likely to find them burdensome, since they have to remember to put them on for the system to work correctly, not to mention the potential privacy issues surrounding collecting detailed information about their movements throughout their waking hours. For this reason, we want to see to what extent [6]'s approach extends to activity discovery using simple binary sensors embedded directly in an environment, without collecting detailed information on people's movements. Most significantly however, the system built by [6] does not have a means to abstract the activities it discovers. As the authors themselves note, people tend to structure and name activities in a hierarchical fashion. Finding a way to automatically detect and output these hierarchies could make activity discovery systems produce more useful and far more semantically meaningful outputs.

3 Our Approach

We now present the approach that we developed for this paper. Our approach is based on that presented by [6], and discussed in Sect. 2. It is built on the concept of topic modelling, a concept taken from the natural language processing community. Due to space constraints, we do not include any detailed introduction to topic modelling here, except to say that topic modelling algorithms take a collection of documents as input, and categorise them into a discrete number of topics based on their respective similarities. Like [6], we utilise an approach to topic modelling called Latent Dirichlet Allocation (LDA). We refer the interested reader to [2] for more detail on LDA.

For our analysis and testing we used the *SCARE corpus* [13], a corpus of situated dialogues, where activities were carried out in a virtual environment by teams of two volunteers. One volunteer controlled movement and action in the virtual environment and one directed this controlling volunteer by providing them with instructions on how to complete a list of given tasks. Since the corpus was collected for research into situated dialogues, we had to convert it into a format more suitable for activity discovery, as in [10]. This was done by processing the full dataset into a list of *events* in the virtual environment. Whenever some aspect of the environment changes (for example, when a door is opened, or when an object is picked up or placed down), we add a new event to our list. The completed list therefore presents the events that took place in the temporal order in which they occurred in the original dataset. Note that only the order is preserved: we do not store information about the time at which the events occurred, or even the amount of time that passed between individual events. The reasons for this will be made clear shortly.

Our specific approach involves the construction of a hierarchy of activities. While this hierarchical process is the primary contribution of this paper, we first discuss the non-hierarchical aspects of the system. Initially, we proceeded to build our system to operate using the same basic algorithm proposed by [6], and discussed in Sect. 2. Figure 1(a) shows an example of a sliding window of length 3 over a dataset of length 5 (in reality, the true SCARE extract we used has a length of over 1600 events). In Fig. 1(b) we can see the sliding window has moved one event forwards. Unlike in the work of [6], we always increment the sliding window by exactly one event at a time, giving our system a larger number of documents to work with. Thus this has produced two documents, and this process continues to generate more documents until the window reaches the end of the dataset. Interestingly, we can also produce a probability vector of topics for each event. This is done by computing the product of the probability distribution vectors for all windows/documents that contain the event, and then re-normalising the resulting vector to produce a probability vector. Distributions produced by this method will be used in Sect. 4 to evaluate the system's performance, and they also provide the basis for the hierarchical analysis.

Our hierarchical analysis aims to address the fact that although a pure topic modelling approach provides a somewhat robust means by which activities can be extracted from raw sensor data, it suffers from a substantial problem. It isn't clear what window size should be used when running the system. This is a user-defined parameter, but setting it to different values can result in profound differences in the resulting output. Worse, the resulting differences may not be simply "right" or "wrong", since as the window lengths increase the level of abstraction of the discovered topics will presumably increase. Many activities could plausibly have multiple levels of abstraction, and an activity discovered at any one level could still qualify as correct. For example, if we were processing a dataset that contained events from a kitchen, we might discover an activity at one level of abstraction that corresponded to "making tea". With a smaller window size, we might then find an overlapping activity corresponding to "boiling kettle" (since boiling kettle could be a constituent activity of making tea). Likewise, with a larger window size we might find an overlapping activity called "making dinner" (again, making tea could be a constituent activity of making dinner).

In order to make progress in solving this issue, our model follows the principal outlined in Fig. 1(c). Here, *Event 2* and *Event 3* were found with high probability to belong to an activity (the probability distribution vector over topics for that event discussed above has a probability exceeding a user-supplied threshold for at least one activity), and the highest probability activity for both matched. This means that we remove them both from the dataset, and replace them with a new event, which abstracts away the activity that this topic is presumed to represent. The algorithm is then rerun for a second iteration using this new dataset, allowing for the inference of higher

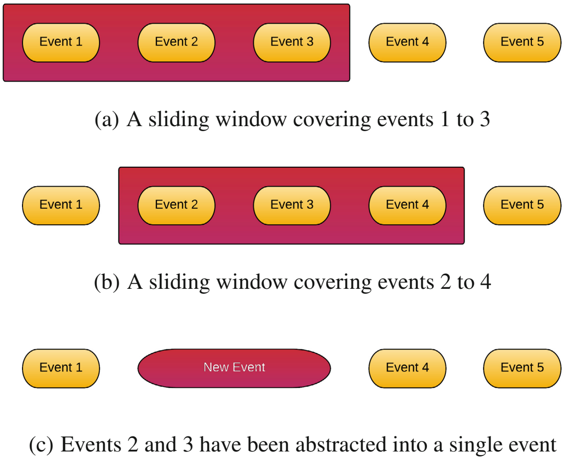


Fig. 1 An illustration of the internal operation of our approach

level activities. This means that by using a fairly small, fixed window length, we can allow the system to explore and discover activities over multiple levels of abstraction. A variant of LDA, Hierarchical LDA (hLDA), exists [1] which already allows for hierarchies of topics to be learned. Our system, however, is doing something conceptually different. Whereas hLDA builds a hierarchy of topics based on perceived semantic similarity (i.e. news about football and basketball could be abstracted into a "*sports*" topic), our system builds hierarchies based on constituency, i.e. where one topic (or activity in our case) is a subset of, or overlaps with, another topic.

4 Results and Discussion

Evaluating the results of an AD system is a challenge. Typically, due to the relationship between pattern discovery in general and clustering, it is assumed that methods for evaluating clustering algorithms (for instance the Silhouette coefficient and similar metrics) would also be appropriate for pattern discovery and thus for AD. However, most of these metrics work by rewarding the maximisation of dissimilarity between cluster members. In many cases, we can imagine distinct activities having a lot of overlap in terms of what sensors they are likely to activate (both using the toilet and taking a shower will activate sensors in the bathroom, but they are clearly distinct activities). This means that these methods may not be suitable for evaluating the performance of AD systems.

We are, however, using a dataset which contains ground truths, which can be compared to the system's output. This is also the performance metric utilised by [6].

Table 1 Performance of our system running with window length 22

Topic	Label	Precision	Recall	F1 score
Topic 0	Move Quad	0.7353	0.5556	0.6329
Topic 1	Move Picture	0.9780	0.1295	0.2288
Topic 2	Move Rebreather	0.8923	0.4947	0.6365
Topic 3	Move Silencer	0.8015	0.1870	0.3032
Topic 4	Move Box	0.0	0.0	0.0
Topic 5	None	0.0449	0.0270	0.0338

We used F1 scores as our evaluation metric. In order for this to happen, we had to find a way of matching topics to their most probable corresponding label in the original dataset. This is done via a simple greedy algorithm. Firstly, we extract the most probable topic/activity for each entry in the dataset according to our algorithm. We then iterate over the ground labels. For each ground label, we evaluate the F1 score for each remaining topic, taking the current label to be the corresponding ground truth. We then assign the topic with the highest corresponding F1 score to be the matching topic for our label. The topic is then removed from consideration before the algorithm repeats on the next label (i.e. we will never assign a topic to more than one label). The results of running this evaluation over a window size of 22 (which seems to perform more acceptably than other sizes we have tried) are shown in Table 4.

At first glance, these results might seem to lag behind that of [6]. However, this is due to a number of important features of this work. To begin with, our dataset is fundamentally more difficult in two ways. Firstly, activities are interleaved in a dataset, and so the boundaries between activities is not as clear-cut as in the dataset used for example by [6]. Secondly, we are using binary sensors that are embedded in the environment, rather than on-body sensors producing complex motion data. This gives access to low-level motions and actions carried out by the user, producing a considerably richer dataset than in some other cases.

Considering the results in more detail, we see that the *Move Box* label seems to have been particularly difficult for the system to detect. Most instances of this activity are very short, so it may not be a large enough pattern for the system to pick up on. The fifth topic was assigned to a label called *None*. This does not appear in the original SCARE dataset: there are 5 labels taken from the dataset, and a sixth label that we artificially added, which in effect means no activity was taking place.

Another problem that could potentially impact on the performance of AD systems is that there may not be a clean overlap between the natural patterns in the data (which is what we would expect such a system to output) and the hand-annotated activities in the dataset itself, which we are using as ground truths for this evaluation. This indicates that our evaluation metric may in fact not be particularly well suited to our problem. It also provides us with no way of evaluating the hierarchical system that we have built. An alternative solution may simply be to visualise the patterns that have been detected in some way, and then manually inspect them. Figure 2 illustrates an example of a diagram we have produced from the system's hierarchical output (the complete image is far too wide to reproduce here without loosing an unacceptable degree of detail). This image can be understood as a graph, with time plotted along the x-axis and the layers of the hierarchy along the y-axis. At the bottom of the image, different colours indicate different sensor events occurring over the course of time. As we begin to move up towards the top of the image, we can see sensor events being replaced with new abstract events, corresponding to low-level activities as explained in the previous section. In the centre left of the image, we can see that an abstract event is created at an early stage in the hierarchy, and this event is subsumed into a new, higher level event about half-way up. At the top of the diagram, this new abstraction is in term subsumed into an even higher one. The very right of the image shows a abstract event being created very late in the hierarchy, but never subsumed by anything. The late formation of abstractions is fairly common, and results from the relative probabilities of events that are carried over from previous layers changing due to the replacement of low-level events with abstract events. On top of this, we can see repeating patterns of colour in the bottommost layer. These correspond to the activities that the system is producing.

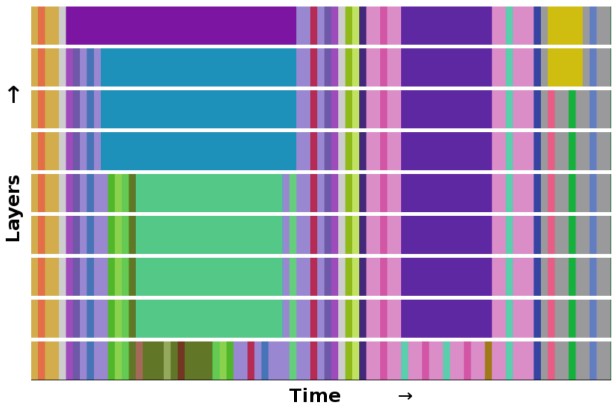


Fig. 2 The diagram style we are using to visualise our system's output

A similarity can be noted between the hierarchical system that we have developed and the concept of maximising compression as the goal of pattern discovery, for example as used by [3]. The concept of using compression in machine learning originates with the minimum description length (MDL) principle, as proposed by [9], and has been used with success before in the AD literature [7]. The topmost layer in the hierarchy is about 86.56% the length of the full dataset in the bottom layer. This means that we have a mathematical justification for claiming that our system works on the SCARE corpus. This approach should also be applicable to other datasets (and we plan to use it on such in the future), but due to time constraints we could not present results for these in this paper.

In conclusion, we have demonstrated and provided an initial evaluation of an algorithm to hierarchically discover activities using an LDA model. We have also presented a reasonable way to visualise the output of this process. As highlighted previously, interleaving remains a significant challenge for activity discovery systems, so we aim to investigate ways of dealing with it in the future. We also hope to look further into evaluation techniques for AD, including finding ways to determine the degree to which the discovered activities make sense semantically to a human observer. Finally, we aim to utilise our approach on larger and more complex datasets.

References

1. Blei, D., Griffiths, T., Jordan, M.: The Nested Chinese Restaurant Process and Bayesian Non-parametric Inference of Topic Hierarchies. Journal of the ACM **57**(2) (2010)
2. Blei, D., Ng, A., Jordan, M.: Latent Dirichlet Allocation. Journal of Machine Learning Research **3**, 993–1022 (2003)
3. Cook, D., Krishnan, N., Rashidi, P.: Activity Discovery and Activity Recognition: A New Partnership. IEEE Transactions on Cybernetics **43**(3), 820–828 (2013)
4. Dong, G., Li, J.: Efficient mining of emerging patterns: discovering trends and differences. In: Proceedings of the fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 1999, pp. 43–52 (1999)
5. Gu, T., Wang, L., Wu, Z., Tao, X., Lu, J.: A Pattern Mining Approach to Sensor-Based Human Activity Recognition. IEEE Transactions on Knowledge and Data Engineering **23**(9), 1359–4347 (2011)
6. Huỳnh, T., Fritz, M., Schiele, B.: Discovery of activity patterns using topic models. In: Proceedings of the 10th International Conference on Ubiquitous Computing, pp. 10–19 (2008)
7. Li, N., Crane, M., Gurrin, C., Ruskin, H.J.: Finding motifs in large personal lifelogs. In: Proceedings of the 7th Augmented Human International Conference 2016, pp. 9–17 (2016)
8. Nguyen, L.T., Tague, P., Zeng, M., Zhang, J.: Superad: supervised activity discovery. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, pp. 1463–1472 (2015)
9. Rissanen, J.: Modeling by shortest data description. Automatica **14**(5), 465–471 (1978)
10. Ross, R., Kelleher, J.: Accuracy and timeliness in ml based activity recognition. In: Proceedings of the AAAI Workshop on Plan, Activity, and Intent Recognition (PAIR), Association for the Advancement of Artificial Intelligence (2013)
11. Ruotsalainen, M., Ala-Keemola, T., Visa, A.: GAIS: a method for detecting interleaved sequential patterns from imperfect data. In: IEEE Symposium on Computational Intelligence and Data Mining, (CIDM 2007), pp. 530–534 (2007)
12. Saives, J., Pianon, C., Faraut, G.: Activity Discovery and Detection of Behavioural Deviations of an Inhabitant From Binary Sensors. IEEE Transactions on Automation Science and Engineering **12**(4), 1211–1224 (2015)
13. Stoia, L., Shockley, D., Byron, D., Fosler-Lussier, E.: SCARE: a situated corpus with annotated referring expressions. In: Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008) (2008)

Estimating the Physical Activity with Smartphones: Analysis of the Device Position and Comparison with GT3X+ Actigraph

Victor H. Rodriguez, Carlos Medrano, Inmaculada Plaza, Cristina Corella, Alberto Abarca and Jose A. Julian

Abstract Nowadays there are commercial devices such as the GT3X+ that can analyze the performance of those who practice some kind of sport, but these devices tend to be rather expensive and complex to use. The objectives of this research are: i) to study the correlation between the measurements of the physical activity with the smartphone and a dedicated accelerometer GT3X+ through the calculation of the counts, ii) to analyze the influence of the position of the smartphone and iii) compare several methods to calculate the energy expenditure trough the counts. Nine volunteers participated in an experiment. They performed different physical activities carrying a smartphone in the right pocket and another one in the hip, together with the GT3X+. The results obtained show a high correlation between the GT3X+ and smartphones for the different types of training (hip and pocket). However the result of the ANOVA indicates that there is no significant difference between the positions of the smartphone.

Keywords Physical Activity · Smartphones · Accelerometer · eHealth

1 Introduction

The lack of physical activity (PA) has a direct influence on cardiovascular diseases and other pathologies such as diabetes mellitus, colon cancer, obesity, hypertension, osteoporosis and depression [1]. In order to have a healthy living, the

V.H. Rodriguez(✉) · C. Medrano · I. Plaza
EduQTech, E.U. Politecnica, c/Atarazana 2, 44003 Teruel, Spain
e-mail: victor_rod_ont@hotmail.com, {ctmedra,inmap}@unizar.es

C. Corella · A. Abarca · J.A. Julian
EFYPAF, Facultad de Ciencias Sociales y Humanas, c/Atarazana 2, 44003 Teruel, Spain
e-mail: {criscor,aabarca,jajulian}@unizar.es

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_6

World Health Organization (WHO) recommends adults aged between 17-64 years perform physical activity for 60 to 150 minutes per week [2].

Ideally, methods for evaluating the PA should be able to quantify energy expenditure during physical activity, as well as the type, duration and intensity of these activities [3]. Traditionally, energy expenditure in humans was determined by the consumption of oxygen and exhalation of carbon dioxide, which is not very practical for everyday environments [4]. Most researchers on PA have focused on inertial sensors such as accelerometers, which can record the movements of the body. These sensors are small and have reduced energy consumption, which makes them ideal for integrating nonobtrusive wearable devices. These features make them suitable for clinical applications [5].

We can find several commercial devices to analyze the performance of people who practice some kind of sport. Their objective is to keep better control over the training. Some of them have been validated. However, most of these devices are based on GT3X, GT3X+ and wGT3-BT actigraphs, which have a considerable cost and in some cases the analysis and recording of data result difficult for medium users. Actigraphs convert signals from the accelerometers in counts, which provide a measure of the physical activity [8] based on the cutoffs proposed by Freedson VM3 2011 [9].

There are several models of actigraphs. They can contain accelerometers, gyroscopes and magnetometers which can be found currently in most commercial Smartphones. Smartphones have become a very popular technology. They can record data continuously in an ubiquitous manner and extract useful information to improve user's wellbeing and health. Some previous studies [10] have shown a strong correlation between records of physical activities taken with smartphones and GT3X+ actigraphs. For that, they made use of the definition of counts based on area under the acceleration curve. In this work we propose an original approach in this field. Specifically, the calculation of the counts from smartphone acceleration signals using a more refined filtering technique. The two contributions of this paper are:

- The calculation of the counts following several models of the literature.
- The study of the influence of smartphone position.
- The study of several methods to measure the energy expenditure.

Counts are a result of summing post-filtered accelerometer values (raw data at 30Hz) into epoch ‘‘chunks.’’ The value of the counts will vary based on the frequency and intensity of the raw acceleration [8].

Therefore, the goal of this paper is twofold:

- To compare the measurement of physical activity with a smartphone and with a dedicated accelerometer, GT3X +, by calculating the counts.
- To analyze the influence of the smartphone position on the results
- To analyze and compare several methods to measure the energy expenditure.

2 Materials and Methods

2.1 Subjects

The subjects for this experiment were nine people between the ages of 25 years and 45 years, in conditions of health sufficiently adequate to be able to perform the activities that they indicate during the tests.

2.2 Experiments

Volunteers were requested to perform common activities. Data collection took 10 minutes per activity. Each record was labelled with the date, time and type of activity. Two smartphones, one in the right pocket and another in the right part of the hip were carried by the participants. At the same time, the GT3X+ device was also placed on the right part of the hip, next to the smartphone (see Fig. 1).

Specifically, the activities performed by the subjects were the following:

- Walking at low speed (pace selected by the volunteer).
- Walking at high speed (pace selected by the volunteer).
- Going up and down stairs.

To record the activities, we used the SAMSUNG GALAXY TREND PLUS GT-S7580 smartphone whose embedded accelerometer is a BOSCH BMC150. The sampling frequency was about 30 Hz. The information obtained was dumped to a text file in the smartphone which was then analyzed using MATLAB. At the same time, the GT3X+ device also recorded the information, which served to validate the samples captured by the smartphone.

Fig. 1 Position of the devices

2.3 Calculating Counts

Method 1 Filtered on Temporal Domain

When the subject is in the rest position, the output of the accelerometer is determined by the orientation of the gravitational vector. On the other hand, when the subject is in movement the accelerometer registers a sample with the acceleration of the motion. You can see an example in Fig. 2.

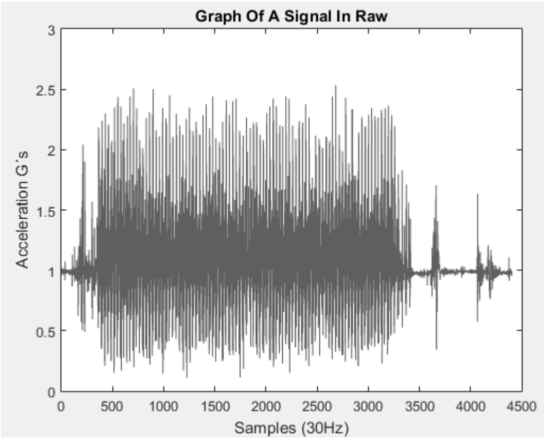


Fig. 2 Motion detection by the accelerometer in G`s

Once an acceleration sample has been registered, the vector magnitude is calculated (Eq. (1)). It contains information about the intensity of the activity carried out by the subject.

$$Vm_i = \sqrt{{x_i}^2 + {y_i}^2 + {z_i}^2} \quad (1)$$

Subsequently, the vector magnitude is computed through a band pass filter, which removes the gravitational component and the high frequency noise (Fig. 3 [9]). The cutoff frequency is the lowest possible, in this case, 0.25 Hz for the high pass and 2.5 Hz for the low-pass [10].

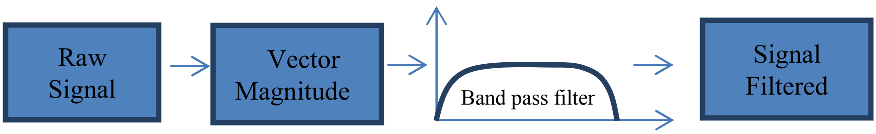


Fig. 3 Signal preprocessing

A recursive filter with infinite impulse response (IIR) of order 2 has been used. The filter coefficients that best fit with the counts provided by the GT3X+ actigraph were obtained by scanning different attenuation values at high and low cutoff frequencies (f_{c1} and f_{c2}) and pass band frequencies (fp).

To convert the values of the vector magnitude into counts, the following steps have been performed:

1. Once the signal is filtered, it is divided into 1 second epochs (30 samples in total, as in the configuration used by the GTX3+ actigraph) resulting in an ray Epc_{ij}, where j represents the epoch, and i is the sample number in that epoch contained in the matrix.

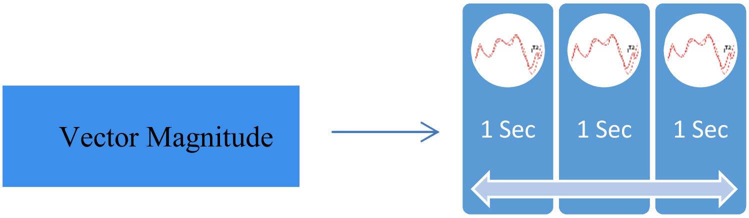


Fig. 4 Vector magnitude fragmented in epochs

2. Taking the 1 second vector magnitude in epochs (EPC), the area under the curve for each vector is calculated (Eq. 2).

$$AUC_j = \frac{1}{L}\sum_{i=1}^{L} |Epc_{ij}| \tag{2}$$

$$Where\ L = size\ of\ the\ window$$
$$i = sample\ between\ the\ epoch$$
$$j = Epoch\ number\ of\ 1\ seg.$$
$$AUC_j = Area\ under\ the\ curve\ for\ the\ j-th\ epoch$$

In [8] it is indicated that, once the area under the curve for each epoch is calculated, the number of counts are obtained by dividing this value by the accelerometer resolution. But as we have no information about the range and resolution of the accelerometer used in the smartphones, this part has been removed. Instead we multiply the counts calculated by the actigraph GT3X+ by 0.0029 (which is the resolution of this device, since its accelerometer has a range of range ±6 g and 12 bits for analog conversion, $0.0029 \approx 12/2^{12}$), thus making the measure independent of accelerometer resolution. Finally, the counts per second are passed to minutes (CPM), as this is the official measure provided by actigraphs.

2.4 Performance Evaluation

The performance estimation was based on the root mean square error (RMSE) and a cross-validation strategy by person-activity, as in [6]. That is, for each run of the cross-validation, data from an activity of a given person was left out (validation set). Then, the filter was optimized with the remaining data (training set). For that purpose, a brute force strategy was conducted. The parameters of the filter were swept (attenuation at f_{c1} from 2 to 4 dB, at f_{c2} from 2 to 4 dB and at fp from 0.5 to 1.9 dB). For each set of parameters the raw signals were filtered and the counts per minute calculated. Then, the value of RMSE was obtained taking as true values the counts obtained from the actigraph. After the swept was completed, the parameters that gave the best RMSE in the training set were selected. In this way, the filter was optimized in that set. Then, the optimized filter was used to process the raw signals in the validation set. The output (counts per minute) was compared with the actigraph to register the value of RMSE in that run of the cross-validation. This was repeated for each person-activity, leading to 27 values of RMSE (corresponding to 3 activities and 9 people).

The training was carried out in three different conditions: taking data only from the pocket, from the hip or both together, while the validation sets were always the same and included data in all the possible positions. An ANOVA analysis with repeated measures was performed to seek for statistical significance between the results of the different training conditions.

3 Results

The results obtained for the different types of training depending on the position of the Smartphone are shown in Table 1. The ANOVA resulted in a $p\ value$ of 0.264, that is, no significant differences in the position of the smartphone were found when the PA is computed with the method used in this paper.

Table 1 RMSE for different types of training

	Pocket	**Hip**	**Pocket + hip**
MeanM1	3.74	3.50	3.49
Std	3.43	2.88	3.00

Fig. 5 shows an example of the counts calculated with the smartphone versus the counts provided by the GT3X+ actigraph. They show a high correlation (Pearson correlation coefficient = .6696).

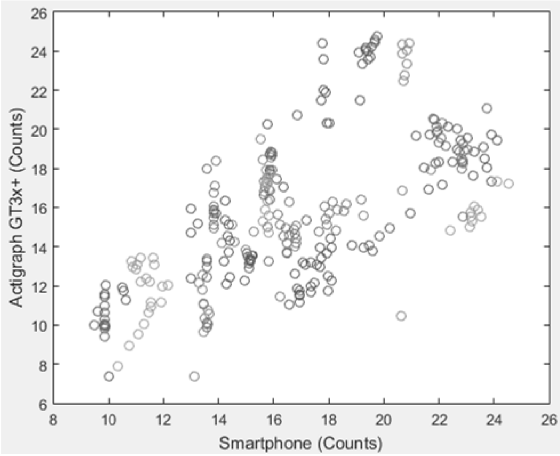


Fig. 5 Smartphone counts vs. GT3X+ actigraph counts

4 Discussion and Conclusions

According to our results, we find that the difference in position between the hip and the pocket does not influence too much the measurement of physical activity with accelerometers of the smartphone. Although the number of activities included in this experiment in still low, this is a reassuring result since smartphones cannot be assumed to be in a fixed position. Thus, using the algorithms described previously can be used to measure the physical activity. On the other hand, the correlation between the counts calculated by the GT3X+ actigraph and those calculated by the smartphone is high, so that the use of smartphones to measure physical activity is quite accurate. This is important since many personal activity recommendations are given in terms of count-based activity levels. However for future work we intend to develop new algorithms to improve correlation by detecting different kind of features and with a more elaborated detection of the kind of activity.

Acknowledgments We would like to thank the "Fondo Social Europeo", the "Diputación General de Aragón" and "Fundación Antonio Gargallo / Fondo Social Ibercaja" for their financial support. V.R. acknowledges a grant from "CONACYT-Gobierno del Estado de Durango, México 330795/386043".

References

1. Liu, S., Gao, R.X., John, D., Staudenmayer, J., Freedson, P.: SVM-Based Multi-Sensor Fusion for Free-Living Physical Activity Assessment. IEEE in Medicine and Biology Society, pp. 3188–3191 (2011)
2. http://apps.who.int/iris/bitstream/10665/44441/1/9789243599977_spa.pdf, recommendations of the World Health Organization (WHO) (accessed January 24, 2016)

3. Liu, S., Gao, R.X., Freedson, P.: Design of a wearable multi-sensor system for physical activity assessment. In: IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics, pp. 254–259 (2010)
4. Mo, L., Liu, S., Gao, R.X., Freedson, P.: Energy-efficient and data synchronized body sensor network for physical activity measurement. In: Instrumentation and Measurement Technology Conference, pp. 1120–1124 (2013)
5. Gyllensten, I.C., Bonomi, A.G.: Identifying Types of Physical Activity With a Single Accelerometer: Evaluating Laboratory-trained Algorithms in Daily Life. IEEE Transactions on Biomedical Engineering **58**(9) (2011)
6. Hekler, E.H., Buman, P.M., Grieco, L., Rosenberger, M., Winter, J.S., Haskell, W., King, C.A.: Validation of Physical Activity Tracking via Android Smartphones Compared to Actigraph Accelerometer: Laboratory-Based and Free-Living Validation Studies. JMIR mHealth uHealth **3**(2), e36 (2015)
7. http://www.actigraph.nl/downloads/getDownload/id/24, ActiLife 6 User's Manual (accessed January 24, 2016)
8. https://help.theactigraph.com/entries/20723176-What-are-counts-, What are counts (accessed January 24, 2016)
9. https://help.theactigraph.com/entries/21452826-What-s-the-difference-among-the-Cut-Points-available-in-ActiLife-, cut points of the actilife (accessed January 24, 2016)
10. Peach, D., Van Hoomissen, J., Callender, H.L.: Exploring the ActiLife® filtration algorithm: converting raw acceleration data to counts. Physiol. Meas. **35**, 2359–2367 (2014). IOP Publishing

An Application for Mobile Devices Focused on Clinical Decision Support: Diabetes Mellitus Case

Lucas Felipe Klein, Sandro José Rigo, Sílvio César Cazella and Ângela Jornada Ben

Abstract Clinical decision-making is performed by health professionals and it is currently connected to the need for manual query for these professionals for clinical guidelines, which are generally formed by large text files, which makes this process very slow and laborious. The development of mobility-based systems, which has the objective of helping the health-related processes, is becoming increasingly widespread. These solutions provide physicians fast and easy access to information anytime and anywhere, in order to support decision-making processes and the quality of care. This paper presents a decision support system based in the Brazilian primary care notebook: Diabetes Mellitus. The solution allows the physicians to interact with a flow of structured questions to assist in clinical reasoning, and, afterwards, they can obtain diagnosis or suggestions of for diabetes treatment. This flow of questions and their outputs are modeled in an ontology, which is found on the knowledge of basic attention specifications. The prototype was tested and evaluated by medical experts, with promising results.

Keywords Mobile devices · Clinical decision support · Primary health care

1 Introduction

Mobile devices are becoming increasingly common in people's lives, since they are in all business areas and help to solve problems. The health sector benefits

---

S. César Cazella(✉) · Â.J. Ben
Federal University of Health Sciences of Porto Alegre, Rio Grande do Sul,
Rua Sarmento Leite 245, Porto Alegre, RS 90050-170, Brazil
e-mail: {silvioc,angelajb}@ufcspa.edu.br

L.F. Klein · S.J. Rigo
Centro de Ciências Exatas e Tecnológicas, University of Vale do Rio dos Sinos,
Av. Unisinos 950, São Leopoldo, RS 93022-000, Brazil
e-mail: lucasfk89@gmail.com, rigo@unisinos.br

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_7

from the use of technology, and the relevance of such devices may have in these areas may be huge. The application of mobile technology in healthcare enables professionals to make their practice more productive, avoiding misunderstanding, and taking a great benefit for both medical staff and for their patients [7].

Among the many potential morbidity diseases, Diabetes Mellitus (DM) is increasing its importance for their expanding prevalence. In South and Central America countries, it is estimated that over 25 million people have diabetes mellitus, and about 27.4% are undiagnosed. Worldwide the numbers are even more alarming, reaching 387 million people with Diabetes Mellitus and 46.3% undiagnosed; thus, one every twelve people in the world present signs of diabetes and almost half of them do not know they suffer from the disease [6]. Diabetes Mellitus is integrated into the primary care sensitive condition, i.e., it is present in a number of health problems that a primary action could be effective in decreasing the risk of hospitalizations. High rates of hospitalization from problems that are present in this set may indicate serious problems of access to health care [1].

The vast amount of information today contributes to the delay in performing diagnostic and clinical analysis, and the difficulty on the part of health professionals in finding certain information may lead often unnecessary hospitalizations [1]. These are usually available in clinical guidelines, which are documents developed with evidence-based statements. It aims to help providers, beneficiaries and other stakeholders about appropriate health interventions. The formalization of these guidelines by computers is called Computer-Interpretable Guidelines (CIG) and makes it possible to develop the decision support that increase the chance of clinical impact compared to using only the written guidelines [9]. To use the guidelines together with the decision support systems can be used an ontology, which serves not only to make the reuse of knowledge easier, but also as a basis for standardization efforts, since they make the concepts explicit, giving terminologies or models. This study presents a prototype for mobile devices developed based on a clinical guideline for Diabetes Mellitus. An ontology is designed to represent the information contained in Guideline Brazilian Clinic for Diabetes Mellitus and also implemented a prototype with defined architecture.

This article is structured into six sections, including the Introduction and conclusion. The second Section contains the theoretical framework used as the basis for the knowledge to develop this work. In Section three, the materials and methods are described. In Section four, the prototype development is presented. Finally, the fifth Section details the evaluation performed.

2 Mobile Health and Diabetes Mellitus

Mobile devices are increasingly present in people's lives, and in health care field it is not different. These applications add value and help a lot in medical activities. For users who work in healthcare devices, they bring big gains for quality and speed of treatment and diagnostics, as well as they are useful either for physicians in training or experts. Also, we may highlight that for ordinary users these devices also bring great benefits, such as heartbeat control while performing physical exercises. For these and other reasons, the union of technology and health is an area that has a bright future. ‘‘The Information Technology can also be used as a potential tool for environmental responsibility policies of different types of business in the health sector, contributing to the rational use of resources and to reduce costs.’’ [13]

In recent years, these two areas are increasingly interconnected. According to the Department of Health of São Paulo [14], ‘‘the technologies are currently an indispensable part of the whole health system. However, in a context where financial resources are limited, the correct development and diffusion of technologies have proven to be a challenge for health systems worldwide.’’

The high cost of diagnosis, treatments, high waste of resourses, among other factors, lead the healthcare sector to suffer to balance their accounts annually. According to [13], reducing costs and increasing efficiency are constant searches, but in the case of hospitals, they are a matter of survival. With the increase of expenses in health, from high-tech use for diagnostics and high waste of resources, the segment finds it difficult to balance their numbers.

The mobile health, or mHealth, came to help to reduce these costs and to improve health as a whole, using the sharp growth of technology aimed at the practice of medicine and public health supported by mobile devices [2]. EHealth is the use of information and communication technologies (ICT) for health. Examples include treating patients, conducting research, health professional education, disease tracking and monitoring of public health [4]. One of the components is the eHealth, mHealth which is defined by the Global Observatory for eHealth (GOE) as medical practice and public health supported by mobile devices such as mobile phones, patient monitoring devices connected to personal digital assistants (PDAs) and other wireless devices [7]. The Global Observatory for eHealth (GOE) is an initiative launched by the World Health Organization (WHO) dedicated to the study of eHealth, its evolution and impact on health of countries [4]. The World Health Organization (WHO) is the directing authority and health coordinator within the United Nations system. It is responsible for providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating options policies based on evidence, providing technical support to countries and monitoring and assessing health trends [10]. Within Brazil there is the Pan-American Health Organization (PAHO / WHO), which, according to [8] "... is an international public health agency with a century of experience, dedicated to improving the health conditions of the countries of the Americas. Integrating the United Nations happens when the entity becomes the Regional Office for the Americas of the World Health Organization. PAHO / WHO is also part of the Organization of American States systems (OAS) and the United Nations (UN)".

Diabetes mellitus is a term that describes a metabolic disorder of multiple etiologies, characterized by chronic hyperglycemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action, or both. The diagnosis of diabetes is increasing its importance due to the fact of increasing prevalence, and the disease is usually associated with dispidemia, hypertension and endothelial dysfunction. It is a health problem that is integrated in Condition Sensitive to Primary, thanks to the evidence showing that the management of this problem in primary care can prevent hospitalizations and possible deaths from cardiovascular and cerebrovascular complications [1]. However, the Diabetes Mellitus may remain asymptomatic for long periods of time and their clinical detection is often made not by symptoms, but by their risk factors. For this reason, it is very important that primary care teams be always vigilant, not just to the symptoms of diabetes, but also to their risk factors [7].

Ontologies are used for knowledge representation. According to [5], it is an explicit specification of a conceptualization. In health-related computing, ontologies have great relevance due to the need for a semantic representation, in which both computers and users may be able to understand. Inside the base, all entities are related and create a structure representing the explicit definition of concepts and their relationships, properties and constraints. According to [3], the entire knowledge base consists of a conceptualization of objects, concepts and entities that are related and have areas of interest. Conceptualization is a simplified and even abstract view of a representation to be created, targeting a specific purpose. Other authors such as [6] cite the ontology definition as an explicit specification of a conceptualization.

3 Methodology

This project uses the knowledge gained to solve a real problem, featuring an applied research. As for his approach, this research is considered quantitative and qualitative. Taking into account the objective of the research, we can conclude that consists of an exploratory research. The research methods used to develop this work were the literature, the survey requirements, case study, prototyping and evaluation of the prototype by a specialist. The survey and the surveys carried out the requirements for this work, related to health care knowledge, with the participation of a specialist in Family and Community Medicine, and Professor of Public Health at the Federal University of Health Sciences of Porto Alegre (UFCSPA).

Referring to the case study, we used the Brazilian guideline for diabetes, more specifically, the ‘‘Care Notebook Basic Diabetes Mellitus in Brazil,’’ published in 2013 by the Ministry of Health. This research has as its main target audience all health professionals. However this work will particularly focus on professional figure responsible for much of the decisions taken in this area, a doctor, specifically physicians who goes by the SUS. One can cite as secondary targets students in the health field who wish to deepen their education, using the proposed solution to train their diagnostic skills and treatment of a clinical guideline. The evaluation was performed by applying a questionnaire to evaluate the prototype which was based on the user experience by three specialists in Primary Health who combat diabetes. Thus, way it was possible to identify the perception obtained by experts on the proposed solution. For data analysis, a quantitative and qualitative approach was used.

4 DMLass's Prototype

This section presents the main points of the analysis, development and application of the proposed system. It also describes the screens and user interaction with the developed prototype. The following are the use cases specified in summary form: a) Perform Diagnostics (UC01): This use case aims to create a consultant patient profile through a previously established questionnaire and use the answers to establish a diagnosis related to Diabetes. Any questions, alternatives and diagnostics are developed based on Ontology; b) Establish Treatment (UC02): the purpose of this use case is to allow the user to establish a treatment for the patient consultant through the profile raised by the questionnaire. Any questions, alternatives and treatments are developed based on Ontology.

The prototype architecture developed was based on the standard Model View Controller (MVC), which organizes the application in three layers, which were the model, view and controller. This pattern were developed in cshtml screens, which are directly linked to a css file for greater control, code reuse and ease of future layout change. These files were placed within this organization in the View layer. Inside the Model layer, classes were developed to be used for the development of future work this application, regarding the administrative area responsible for editing and creating ontologies. In Control layer, they were created the control files and access to services, where the requests, welcome the views are processed and the structure of views are determined through information from the access to ontology. Access to ontology, from the service, is made by dotNetRDF library, which is an open source tool whose main objective is to provide an effective way to work with RDF files within a .Net application.

The ontology has an architecture with two main aspects, the concepts of the addressed area of knowledge, such as patient, disease, symptom, treatment, etc. and also the concepts of the process used for the diagnosis and prompt treatment by the prototype. Both parts were based on Brazilian primary care notebook: Diabetes Mellitus, but they lack integration. This will be developed in future projects and used to improve the usability of the prototype. The developed ontology has two main goals: 1) the representation of the main concepts related to the attention area of the created model, which involves elements such as ‘‘patient’’, ‘‘symptom’’, ‘‘treatment’’, among many others; 2) the representation of knowledge that integrates the experience and practice of physicians in the use of clinical guidelines as a support tool and guide to the procedures for diagnosis and treatment. For the representation of concepts needed it was used as a basis text of medical clinical guideline on diabetes. The results can be seen partially in Figure 1.

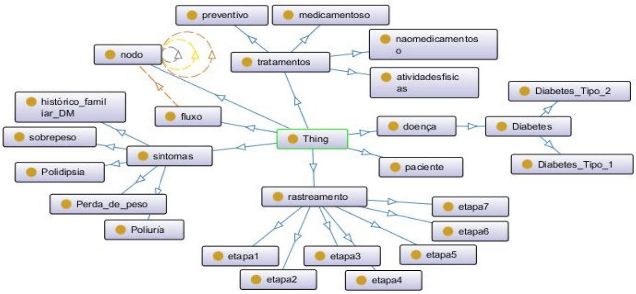
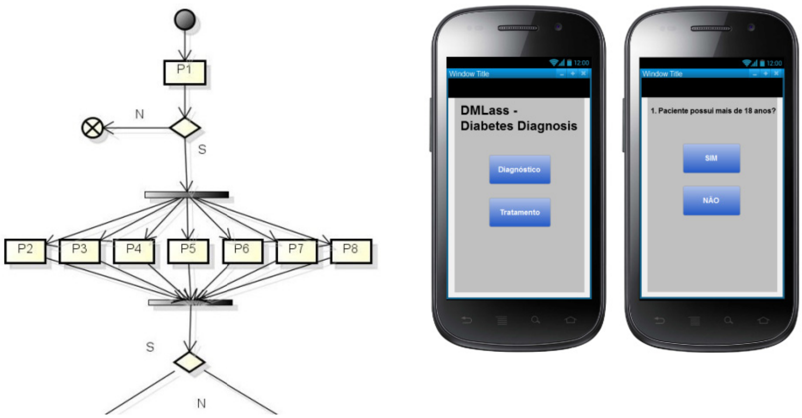


Fig. 1 Ontology View

For representing the activities and steps involved in diagnosis and treatment, experts were consulted primary care area in diabetes and family doctors and community. The description of experts was modeled as two flows elements (diagnosis and therapy). These elements gave rise to the equivalent concepts in the ontology, which can be consulted in accordance to the need for information to support activities.

Figure 2 (a) represents a small part of the diagnostic flow where the black dot represents the beginning of the stream, which is fired from the diagnostic click the button from the home screen. P1 is the first question, which has two possible answers, where ‘‘N’’ response (NO) leads to an exit and ‘‘Y’’ (YES) leads to a set of questions. This group of questions was developed because if any of the questions get the answer ‘‘S’’ (YES) the flow must follow until P9; in case of answer ‘‘N’’ (NOT) the flow goes to the next question of the group. If all group answers are "N" (NOT) the flow goes to another group of questions.



(a) (b)

Fig. 2 (a) Diagnostic Flow and (b)Prototype's Interfaces

The prototype has two interfaces that vary according to the selected option. An example of interfaces can see in Figure 2 (b). The ontology has two streams, diagnosis and treatment, both have interconnected questions and in some cases outputs, representing a diagnosis or a treatment. Construction of the prototype screens search the first question and presents to the user and according to the response, the system searches the next question, to find a way out. The first interface shown in Figure 2 (b), is the initial screen of the prototype, where the user can choose whether to make the diagnosis or treatment of the patient. The treatment option is available on the home screen for the diagnosis of the patient may already have been done at another time and treatment can now be started. The second interface shown in Figure 2 (b) is the first question of the diagnosis, which was returned by this ontology. The response options vary depending on the question, and the ontology are created and linked stages of the questions flow.

5 Evaluation and Results

The completion of this assessment was based on a questionnaire used as data collection instrument. From the answers obtained the usability of the prototype can evaluated, ie.. the degree that the user believes that the application supports their efforts. In the background, the assessment is to evaluate the basics of the Diabetes Mellitus provided by the system. Therefore, this data collection has to evaluate the proposed tool for solution of the problem as well as to see if it reaches the proposed objectives initially. The questionnaire consists of seventeen issues, which are divided into three groups. The first group is formed by questions from one to six and has the purpose of collecting data on the respondent's profile and their familiarity with the use of decision support systems. This group consists of four free-choice questions and two free responses. The second group of questions has aimed to evaluate the acceptance of the prototype by the respondents, and it is composed of seven to fourteen issues, as follows: (a) six questions using a Likert scale of five points, where the value ‘‘5’’ means ‘‘Agree’’ and ‘‘1’’ means ‘‘Strongly Disagree’’; (B) a free choice issue; and (c) a matter of free response. Finally, the last group consists of the questions from fifteen to seventeen, all essay intended to comments and suggestions for improvement of the system and of the ‘‘Guide to Diabetes Mellitus’’.

In carrying out this evaluation, medical healthcare specialists, some of which are also university professors, were invited. The evaluation was answered by three experts, who are part of the target audience of this research, i.e., family doctors and community. After using the prototype, the evaluators answered the questionnaire developed and sent it online and anonymously. All invited experts agreed to participate in the evaluation process, reaching a total of three evaluators. The profile of the respondent sample was composed primarily by health professionals, two of which are also university professors. Their main practice area is family and community, and they have around 2-10 years of experience in the area of operation. Also, in the first group of questions, the questionnaire has a question of familiarity of respondents with the use of decision support systems. Only one of the respondents is familiar with this type of system.

The second part of the questionnaire was formed by questions 7 to 14, which involved the evaluation of the acceptance of the prototype by the respondents. The answers to questions 7, 8 and 9 showed that 67% of respondents agreed that usability and system features are easy to understand (question 7), to use (question 8) and that the features are clear and objective (question 9). As for question 10, varied opinions were obtained regarding the possibility of the application developed to replace the use of printed clinical guideline (part agreed and some disagreed). In question 11, it was asked about the possibility of inclusion and change content directly in the system developed without the assistance of a professional information technology, allowing, through these changes, the greater consistency of information, and largely agreed how much would be appropriate to that function.

Finishing the second group, we have questions 13 and 14. The question 13 checks if the respondents would use this system developed for patient care. To help them to better understand the previous answer, the question 14 asked the justification for the use, or otherwise, of the application. Analyzing the reasons given by the respondents, it is clear that the application brings some benefits – according to the respondents, thanks to its practicality and ease of use. There was only one case of denial about the use of the application, but the respondent's justification says that it would not use because their knowledge is extensive in this area.

Questions 15 and 16, which make up the last group of questions in the questionnaire were intended to give space to comments and suggestions for improvements, both for the system (issue 15), and for the ‘‘Guide to Diabetes Mellitus’’ which was developed during this research (question 16). Even with a good margin to its acceptance among responders, it can be seen that some further improvements in the developed system can and should be made.

6 Conclusions

This research propose a to facilitate and perhaps enhance, the process of clinical decision-making through a system based on case study of the Brazilian primary care notebook: Diabetes Mellitus. By using a centralized database hosted by the site, all the additions, changes or deletions to the contents of Ontology, will be automatically and instantly updated for all users with only the use of the page refresh the browser accessed. This fact proves very important because of the need for professionals in the field of health are always updated, in order to obtain a diagnosis and precise treatment, favoring the quality of care by physicians regarding their patients.

After performing all of the steps used in the development of this system, and especially acceptance that it has obtained from the target public, it is believed that the goal previously established for this research has been reached. We also understand that the proposed solution, is directly connected to an existing need in health, because its features were able to speed up, facilitate or even enhance the decision-making process, providing the highest quality medical care to their patients.

References

1. Alfradique, M.E., et al.: Internações por condições sensíveis à atenção primária: a construção da lista brasileira como ferramenta para medir o desempenho do sistema de saúde (Projeto ICSAP – Brasil). Cadernos de Saúde Pública, Rio de Janeiro **25**(6) (2009)
2. Cazella, S.C., Feyh, R., Ben, A.J.: A decision support system for medical mobile devices based on clinical guidelines for tuberculosis. In: Ramos, C., Novais, P., Nihan, C.E. (Org.) Advances in Intelligent Systems and Computing, 1st edn., pp. 217–224. Springer International Publishing (2014)
3. Genesereth, M.R., Nilsson, N.J.: Logical foundations of artificial intelligence. Morgan Kaufmann Publishers Inc., San Francisco (1987)
4. GOE. Global Observatory for eHealth (2014). http://www.who.int/goe/en/ (accessed: November 28, 2014)
5. Gruber, T.R.: A translation approach to portable ontology specifications. Knowl. Acquis. **5**(2), 199–220 (1993)
6. International Diabetes Federation. Diabetes atlas update 2014: Regional & Country Facctsheets. http://www.idf.org/sites/default/files/Atlas-poster-2014_EN.pdf (accessed: maio 2, 2015)
7. World Health Organization. MHealth: New horizons for health through mobile technologies, Global Observatory for eHealth series – V. 3 (2011)
8. PAHO, Organização Pan-Americana de Saúde. http://www.paho.org/bra/index.php?option=com_content&view=article&id=885&Itemid=672 (accessed: November 28, 2014)
9. Peleg, M.: Computer-interpretable clinical guidelines: A methodological review. Journal of Biomedical Informatics **46**(4), 744–763 (2013)
10. WHO, World Health Organization (2014). http://www.who.int/about/en/ (accessed: November 28, 2014)

Customized Normalization Method to Enhance the Clustering Process of Consumption Profiles

Catarina Ribeiro, Tiago Pinto and Zita Vale

Abstract The restructuring of electricity markets brought many changes to markets operation. To overcome these new challenges, the study of electricity markets operation has been gaining an increasing importance.With the emergence of microgrids and smart grids, new business models able to cope with new opportunities are being developed. New types of players are also emerging, allowing aggregating a diversity of entities, e.g. generation, storage, electric vehicles, and consumers. The virtual power player (VPP) facilitates their participation in the electricity markets and provides a set of new services promoting generation and consumption efficiency, while improving players` benefits. The contribution of this paper is a customized normalization method that supports a clustering methodology for the remuneration and tariffs definition from VPPs. To implement fair and strategic remuneration and tariff methodologies, this model uses a clustering algorithm, applied on normalized load values, which creates sub-groups of data according to their correlations. The clustering process is evaluated so that the number of data sub-groups that brings the most added value for the decision making process is found, according to players characteristics. The proposed clustering methodology has been tested in a real distribution network with 30 bus, including residential and commercial consumers, photovoltaic generation and storage.

Keywords Clustering · Consumption profiles · Dynamic tariffs

C. Ribeiro · T. Pinto(✉) · Z. Vale
GECAD – Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development, Institute of Engineering, Polytechnic of Porto (ISEP/IPP), Porto, Portugal
e-mail: {acrib,tmcfp,zav}@isep.ipp.pt

The present work was done and funded in the scope of the following projects: People Programme of the European Union's Seventh Framework Programme FP7/2007-2013/ project ELECON, REA grant agreement No 318912; EUREKA - ITEA2 Project M2MGrids with project number 13011.

© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_8

1 Introduction

The restructuring of the electricity sector is characterized by an enormous increase in competition and profound changes in the participant entities. The restructuring was performed so that the competitiveness could be increased, but it also had exponential implications in markets complexity and unpredictability [1]. Potential benefits will depend on the efficient operation in the market and, on the other hand, in the remuneration of aggregated players. Important developments concerning electricity market players modelling and simulation including decision-support capabilities can be widely found in the literature [2-3].

Much like electricity markets, subsystems of the main network are rapidly evolving into a reality, coordinating these entities is a huge challenge that requires the implementation of distributed intelligence, potentiating the concept of Smart Grid (SG) [4, 5]. However, the two concepts are not converging towards common goals and technical and economic relationships are addressed in an over simplistic way. Present operation methods and Electricity markets models do not take full advantage of installed DG, yielding to inefficient resource management that should be overcome by adequate optimization methods [6]. Player aggregating strategies allows players gaining technical and commercial advantages, individuals can achieve higher profits due to specific advantages of a mix of technologies to overcome disadvantages of some technologies. The aggregation of players gives rise to the concept of Virtual Power Player (VPP) [7]. VPPs aggregate different types of resources. Each aggregated player has its individual goals; VPPs should conciliate all players in a common strategy able to enable each player to pursuit its own objectives [8].

This paper proposes a data mining methodology, based on the application of a clustering process, which groups the typical load profile of the consumers of a Smart Grid according to their similarity. The separation of consumers in different groups allows proposing specific consumption tariffs to each group, so that consumers' load profile is taken into account to meet the objectives of the Smart grid aggregator. This methodology is tested using a smart grid that includes several real consumers of different types (residential and commerce). This work thus proposes a customized normalization method to treat data before it is used by the clustering process.

2 Smart Grid and Electricity Markets Simulation

Many works have been developed using simulators to model the complex interaction between electricity market players. Successful examples sustain the fact that a multi-agent system (MAS) with adequate simulation abilities, is the best approach for simulating electricity markets [9-11]. The Multi-Agent Simulator for Competitive Markets (MASCEM) is a platform that simulates several electricity market types, while providing decision support to players' actions [9]. This type of simulators are able to represent market mechanisms and players' interactions.

However, for them to be valuable decision support tools in foreseeing market behaviour, they need to be used in testing adequate and realistic scenarios. Real data analysis by means of a knowledge discovery process will be a crucial step forward to assure that MASCEM agents exhibit adequate profiles and strategies.

Multi-Agent Smart Grid simulation Platform (MASGriP) [10] simulates, manages and controls the most important players acting in a Smart Grid environment. This system includes simulated players, which interact with agents that control real hardware. The considered players include operators, and energy resources, such as several types of consumers, producers, electric vehicles, among other. Aggregators are also considered, namely: VPPs and Curtailment Service Providers (CSP) [12]. These players introduce a higher level of complexity to the management of the system. Joint simulations of MASCEM and MASGriP enable a simulation environment that includes the participation of Smart Grid players in electricity markets, or even internal Smart Grid markets, using complex markets models provided by MASCEM.

3 Decision Support Tool for Electricity Markets Remuneration and Tariff Definition

The Remuneration and Tariff Mechanism (RemT) [13-14] is a decision support mechanism that is being developed to support the VPP actions in the definition of the best tariff and remuneration to apply to each of the aggregated players, regarding the VPP objectives and the individual goal of each aggregated player. VPPs in the scope of MASCEM use RemT to remunerate aggregated players, according to the results obtained in the electricity market, the penalties for breach of contract, contracts established to guarantee reserve, demand response programs and incomes of aggregated consumers. The definition of remuneration and tariffs is based on the identification of players' types and on the development of contract models for each player type. This considers players with a diversity of resources and requirements, playing several distinct roles (a player can be a consumer, a producer and can be responsible for one or several V2G). The terms for new contracts and the best strategies for each context are determined by means of machine learning methods and data-mining algorithms.

3.1 Clustering Approach

A wide variety of clustering algorithms can be found in the literature and unfortunately, there is no single algorithm that can by itself, discover all sorts of cluster shapes and structure [15]. K-means [16], has been used, as it proves to be a robust model for distinct applications: K-means minimizes the distance from each point to the centre of the respective cluster, as defined in (1).

$$\min \sum_{i=1}^{k} \sum_{x \in c_i} \left\| x - \mu_i \right\|^2 \tag{1}$$

where μ_i is the mean of points in C_i, *i.e.* the cluster *centroid.* To determine the quality of the division of players into different clusters the clusters validity indices MIA and CDI [17] have been used, as formalized in (2) and (3) respectively.

$$MIA = \sqrt{\frac{1}{K} \sum_{k=1}^{K} d^2(x^{(k)}, \mu^{(k)})} \tag{2}$$

$$CDI = \frac{\sqrt{\frac{1}{K} \sum_{k=1}^{K} \left[\frac{1}{2.n^{(k)}} \sum_{n=1}^{n^{(k)}} d^2(x^{(m)}, \mu^{(k)}) \right]}}{\sqrt{\frac{1}{2K} \sum_{k=1}^{K} d^2(x^{(k)}, R)}} \tag{3}$$

where d represents the Euclidian distance between two points, and R is the representative load profile of all consumers.

3.2 Customized Normalization

Analysing the results of previous work [13], is possible to conclude that aggregation strategies have very good results, and are very useful, because they provide a good separation according to what is intended. The non-normalization grouping process has led to a clear separation between different consumers types, as it considers the absolute consumption amounts in the clustering process. The normalized data, used as formalized in (4) and (5), reveals a separation through consumption profiles, although it is not able to consider the differences in consumption quantity.

$$N_{c,h} = \frac{L_{c,h}}{ML_c}, \forall c \in co \tag{4}$$

$$ML_c = \max(L_c), \forall c \in co \tag{5}$$

where N is the common normalized load, for each consumer c, for each hour h, and co is the set of all considered consumers. ML is the largest consumption value, of the consumer c, considering all hours.

To improve the results achieved in the previous works, the customized normalization process is introduced. This method normalizes data using each consumers' load value at each period divided by the largest recorded value of all loads in all periods, it is formalized in (6) and (7).

$$SN_{c,h} = \frac{L_{c,h}}{SML_h}, \forall c \in co \tag{6}$$

$$SML_h = \max(L_{co,h}), \forall c \in co \tag{7}$$

where *SN* is the load with a different normalization process, for each consumer *c*, for each hour *h*. *SML* is the largest consumption value recorded for all consumers at the time *h*.

The proposed customized normalization method aims to combine the advantages of both previous approaches (using non-normalized data, and regular normalization), so as to achieve consumer groups that capture both differences in the quantities of consumption, and also the trends of consumer profiles along the hours. The clustering process takes into account the tendency of the consumption values trough the time, regardless of its absolute amount. This separation is very important, according to different consumers′ types and profiles, it works as a base for personalized and dynamic consumption tariff definition. Using this approach ensures that the data are also normalized in a range between 0 and 1, but without losing information related to differences between amounts of consumption among consumers. While using the regular normalization, the value 1 is attributed to the greater consumption value of each consumer (thus both consumers with large and small values will always have one value of 1 in a certain hour), using the customized normalization method, only the largest consumer of all, will have a value of 1. The smaller consumers will have normalized values with smaller values, proportional to the difference between the quantities of consumption of that consumer and the largest consumer in each hour. Thus normalization is still made between 0 and 1, but there is visible difference between higher and lower consumption among different consumers, and the evolution of consumption of each consumer profile is also captured.

4 Case Study

This case study intends to show the adequacy of the proposed customized normalization clustering methodology to solve the problem of remuneration of players with heterogeneous characteristics and behaviors. In order to test the adequacy of the method, a clustering algorithm has been applied, concerning the consumption data of a total of 82 consumers (8 residential houses, 8 residential buildings with 72 loads, and 2 commercial buildings). Data has been collected from a real distribution network throughout one year. The Smart grid accommodates distributed generation (photovoltaic and wind based generation) and storage units, which are integrated in the consumption buildings. The accommodated photovoltaic generation, wind based generation and storage units are related to the building installed consumption power, according to the current legislation in Portugal. Further details on the considered distributed network can be seen in [18].

The K-means algorithm has been used to perform the clustering process using non-normalized values of load (section A), and also normalized values, using both the regular normalization method (section B) and the proposed customized normalization method (section C).

4.1 Non-normalized Data

The clustering process is performed for different numbers of clusters, in order to enable grouping consumers according to the similarity of their consumption profiles, in order to support the definition of specific tariffs that are suited for each of the consumer groups. From [13] it has been concluded that, by analyzing MIA and CDI results from the clustering of non-normalized data, the best clustering results are achieved with the use of 3 clusters, as the clustering error is minimal. When using 2 clusters, a clear separation of residential houses and buildings from commercial buildings is visible. It is also visible that the two commercial buildings (corresponding to loads 1 and 2) have been allocated to cluster 1, and the rest of the loads, corresponding to residential consumers, have been aggregated in cluster 2. This can be observed in Figure 1 which presents the load profiles of consumers that have been grouped in cluster 1 and in cluster 2 using the non-normalized data.

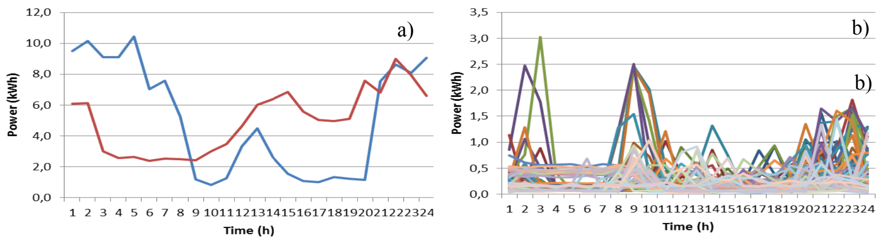


Fig. 1 Consumption profile of loads allocated to: a) cluster 1; b) cluster 2

From Figure 1 it is visible that cluster 1 includes the two commercial buildings, with very distinct load profiles, and cluster 2 includes all the residential buildings and houses. When considering the grouping process with 3 clusters, the difference is that there is still a separation from residential houses and buildings to the commerce. However, in this case the two types of commercial buildings are also separated, as they present very different load profiles.

4.2 Normalized Data

In the second clustering process, regular normalized data were used. The normalization was made considering each type of consumer. The value of load corresponding to each period was divided for the maximum value register in that specific load in the 24 periods. When using normalized values, a more accentuated descent of the clustering error values is visible. The descent in the error value is however, stable from the start, which hardens the identification of the optimal number of clusters that should be used. For this reason it is not advantageous to use more than 2 or 3 clusters, since the use a larger number is not reflected by a significant gain in clustering error. By analysing the results of the clustering process with 2 clusters, it can be seen that the separation is not as clear as it was with non-normalized values. The two commercial buildings corresponding to load 1 and 2, were aggregated in different clusters, together with several residential consumers. However, the clustering process with normalized values has better results from the load profile separation stand point, as can be seen from Figure 2, which presents the allocation of the consumers to the different clusters, when considering normalized data and 2 clusters.

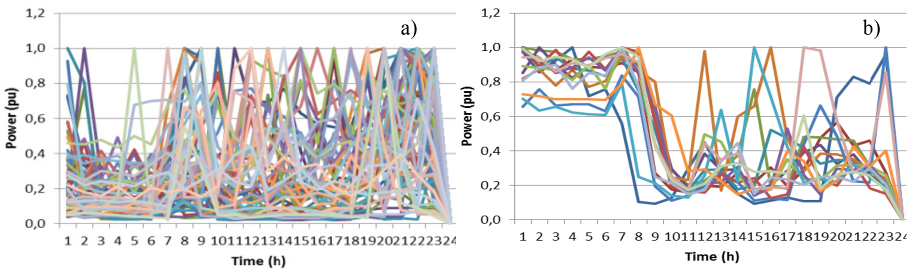


Fig. 2 Consumption profile of loads represented in: a) cluster 1, b) cluster 2

From Figure 2 it is visible that although the consumer types cannot be separated correctly with this approach as occurs when using non-normalized data (Figure 1), the separation of the load profiles is more evident in this case, since profiles are grouped independently from the gross amount of consumption itself.

4.3 Customized Normalization

The clustering process is performed for different numbers of clusters (from 2 to 6). MIA and CDI are used to analyse the clustering error. Figure 3 presents the comparison of the MIA and CDI error values that are achieved when using from 2 to 6 clusters, with each of the three considered methods: non-normalized data, normalized data using the regular method, and customized normalization method.

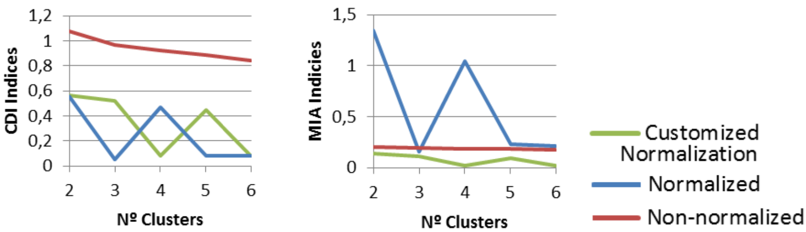


Fig. 3 CDI and MIA results for different numbers of clusters, using non-normalized, normalized and customized normalization

In Figure 3 it is possible to see that the proposed method has a lower error when compared with the previous methods. The best clustering results are achieved with the use of 4 clusters, as the clustering error is minimal. With the use of 4 clusters, the two commercial buildings are separated into a different cluster each, the third cluster allocates some of the residential buildings that have similar load profiles, and the rest of the loads, corresponding to residential houses and some residential buildings, have been aggregated in the final cluster. Figure 4 represents the load profile of the different consumer types considering 3 clusters. Figure 4 a) represents the consumption profiles of the two commercial buildings, which, as it is possible to see, have very different consumption profiles, especially during the night. This is why they were allocated into different clusters when 4 clusters are considered. Figure 4 b) and c) represent the consumption profiles of the loads allocated to the other two clusters.

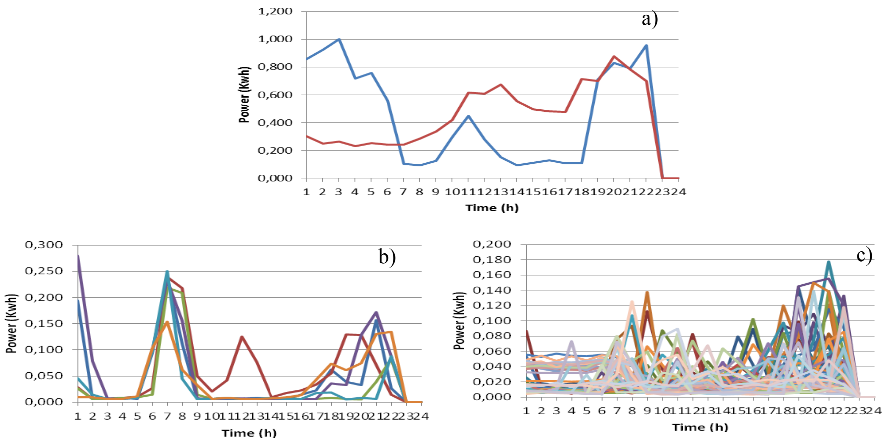


Fig. 4 Consumption profile of: a) commercial buildings, b) cluster 2; c) cluster 3

From Figure 4 the separation of the load profiles is evident. It is also visible that a separation taking into account the gross amount of consumption of each consumer has been accomplished, as commercial consumers, which present much higher consumption values, have been separated from the residential consumers.

The proposed customized normalization brings, therefore, clear advantages to the RemT tariff definition process. It enables to clearly identify different consumers, taking into account their consumption tendency and amount, therefore breaking the way for an objective and fare definition of dynamic electricity tariffs, which can suitably fit each of the identified groups, i.e. consumers with similar consumption tendencies, taking into account their dimension. The new type of normalization, when compared with the previous normalization types, allows an even more clear separation of consumer types, which is evident from the load profile graphs that show the separation into different clusters, and also by the MIA and CDI values, which show that the proposed method achieves smaller clustering error values than the other methods.

5 Conclusions

Electricity markets are experiencing profound transformations. Currently there is a gap in what concerns VPP aggregated players' tariffs and remuneration. In order to overcome this problem, developing appropriate methods is essential. This paper presents the development of a tool that provides a decision support for VPP definition of best tariff and remuneration to apply to each aggregated player, RemT. To develop RemT a clustering methodology that uses different data normalization methods was presented, and a new customized normalization method has been introduced.

The results of the presented case study, based on real consumption data, show that the customized normalization method combines the advantages of both previous approaches, so as to achieve more consumer groups that capture both differences in the quantities of consumption, as well as the trends of consumer profiles along hours. This is crucial, according to different consumers´ types and profiles, as it works as a basis for personalized and dynamic consumption tariff definition. Thus normalization is the same made between 0 and 1, but there is visible difference between higher and lower consumption among different consumers, and the evolution of consumption of each consumer profile is also captured. RemT mechanism is evolving to become a crucial tool to go a step forward in electricity markets simulation, by enabling a fair and dynamic means to define electricity tariffs for different types of consumers.

References

1. Meeus, L., et al.: Development of the Internal Electricity Market in Europe. The Electricity Journal **18**(6), 25–35 (2005)
2. Praça, I., Ramos, C., Vale, Z., Cordeiro, M.: MASCEM: A Multi-Agent System that Simulates Competitive Electricity Markets. IEEE Intelligent Systems **18**(6), 54–60 (2003)
3. Koritarov, V.: Real-World Market Representation with Agents: Modeling the Electricity Market as a Complex Adaptive System with an Agent-Based Approach. IEEE Power & Energy magazine, 39–46 (2004)
4. Shahidehpour, M., et al.: Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management. Wiley-IEEE Press, pp. 233–274 (2002)
5. Blumsack, S., Fernandez, A.: Ready or not, here comes the smart grid! Energy **37**(1), 61–68 (2012)
6. Sousa, T., et al.: "Intelligent Energy Resource Management Considering Vehicle-to-Grid: A Simulated Annealing Approach. IEEE Trans. on Smart Grid **3**, 535–542 (2012)
7. Vale, Z., Pinto, T., Praça, I., Morais, H.: MASCEM - Electricity markets simulation with strategically acting players. IEEE Intelligent Systems **26**(2) (2011). Special Issue on AI in Power Systems and Energy Markets
8. Pinto, T., et al.: Multi-agent based electricity market simulator with VPP: conceptual and implementation issues. In: 2009 IEEE PES General Meeting (2009)

9. Pinto, T., et al.: Strategic bidding methodology for electricity markets using adaptive learning. In: Modern Approaches in Applied Intelligence, vol. 6704, pp. 490–500 (2011)
10. Oliveira, P., et al.: MASGriP - a multi-agent smart grid simulation plataform. In: IEEE Power and Energy Society General Meeting, San Diego, California, USA, pp. 1–10 (2012)
11. Vale, Z., et al.: Mascem - electricity markets simulation with strategic agents. IEEE Intelligent Systems, Special Issue on Agents and Markets **26**(2), 9–17 (2011)
12. Kieny, C., et al.: On the concept and the interest of virtual power plant: some results from the european project FENIX. In: 2009 IEEE Power and Energy Society General Meeting, Calgary, Canada, July 26–30, 2009
13. Ribeiro, C., et al.: Data mining approach for decision support in real data based smart grid scenario. In: IATEM (2015)
14. Ribeiro, C., et al.: Intelligent remuneration and tariffs in for virtual power players. In: IEEE PowerTech (POWERTECH) Grenoble, France, June 16–20, 2013
15. Jain, A.K., et al.: Data Clustering: A Review. ACM Computing Surveys **31**(3), 264–323 (1999)
16. Jain, A.K.: Data Clustering: 50 years beyond K-Means. Pattern Recognition Letters **31**(8), 651–666 (2010). Elsevier
17. Chicco, et al.: Support Vector Clustering of Electrical Load Pattern Data. IEEE Transactions on Power Systems **24**(3), 1619–1628 (2009)
18. Canizes, B., et. al.: Resource scheduling in residential microgrids considering energy selling to external players. In: Power Systems Conference (PSC 2015), South Carolina, USA, March 10–13, 2015

Dynamic Traffic Light Control System Based on Process Synchronization Among Connected Vehicles

Khac-Hoai Nam Bui, O-Joun Lee, Jason J. Jung and David Camacho

Abstract Vehicular traffic is tremendously increasing around the world, especially in large urban areas. The resulting congestion has become a key issue and emerging research topic to transportation specialist and decision makers. In this study, inspired by recent advanced vehicle technologies, we take into account in improving traffic flow in real-time problem. In order to solve the problem, we propose a new approach to manage traffic flow at the intersection in real-time via controlling by traffic light scheduling. In particular, the proposed method is based on process synchronization theory and connected vehicle technology where each vehicle is able to communicate with others. The traffic deadlock is also taken into consideration in case of high traffic volume. The simulation shows the potential results comparing with the existing traffic management system.

Keywords Connected vehicles · Process synchronization · Intelligent transportation systems · Real-time processing · Intersection management

1 Introduction

Nowadays, it is obvious that the volume of traffics is rapidly growing on roads worldwide. This result affects directly to economies, human health, and environment because of huge traffic congestion. By this way, Intelligent Transportation Systems (ITS) have been developed to deal with these emergent problems. ITS have been regarded as advanced applications which applied to transport and infrastructure to

K.-H.N. Bui · O-J. Lee · J.J. Jung(✉)
Department of Computer Science and Engineering, Chung-Ang University, Seoul 156-756, Korea
e-mail: hoainam.bk2012@gmail.com, concerto9203@gmail.com, j2jung@gmail.com

D. Camacho
Computer Science Department, Universidad Autóma de Madrid, Madrid, Spain
e-mail: david.camacho@uam.es

© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_9

transfer information between systems for improved safety, productivity and environmental performance [1]. Among many kinds of applications of ITS, traffic control at intersections is always a significant challenge for the researchers and the developers, because of their complex locations and various shapes.

There have been many approaches to deal with the traffic congestion at intersection. Traffic light scheduling is the traditional approach where the vehicles stop-and-go according to the control of traffic light. The state-of-the-art issues in this approach focus on adaptive and dynamic traffic light system. There have some papers using optimized signal solutions to solve this problem. Chen and Shi [2] proposed a Genetic Algorithm for optimization the signal at intersection. On the other hand, the approaches [3, 4] used fuzzy logic for traffic signal control. Srinivasan et al. [5] introduced the Neural Network to deal with real-time management. Moreover, some approaches focus on optimizing time of green light [6, 7]. These approaches are interesting and come out with good result, however, they are hard to implement since the complexity of computation. To deal with this problem, Wu et al. [8] introduced a new approach for dynamic the traffic light control based on distributed mutual exclusion algorithm. Although the system will be complicated in case of high traffic flow. With the developing of internet of things, many studies suggested the use of Wireless Sensor Network (WSN) [9, 10, 11, 12] for adaptive and cost-effective traffic control. However, with designing large number of sensor, the energy-efficient problem becomes a big challenge for the developers. In this regards, using Vehicular Ad hoc Network (VANET) is a good solution for energy problem. Bento et al. [13] proposed the intelligent traffic managements based on VANET. By this way, the vehicles are able to avoid the conflict based on send their signal. Nevertheless, the vehicles are only able to communicate in short range of transmission.

Recently, the 'connected vehicles' concept has been identified as the most promising future solution to enhance road traffic conditions and achieve the goal for more efficient and sustainable traffic management solutions. It refers to the wireless connectivity enabled vehicles that can communicate with their internal and external environments such as supporting the interactions of V2S (vehicle-to-sensor), V2V (vehicle-to-vehicle) and V2I (vehicle-to-internet) [14]. In these interactions, vehicles are enhanced the situational awareness not only collecting the information about themselves but also communicating with other objects such as traffic management system and others vehicles. Based on these advanced technologies, in this work, we propose a new solution for controlling traffic flow at intersection. In particular, each vehicle pass into the intersection is regards as a process which executing in critical section (core area of the intersection). In this regards, how to manage the traffic flow if there have many processes want to execute in the common resource (race condition) as well as improving the waiting time of each process in case of high traffic flow are the main challenges that this study focus on. Moreover, the deadlock can occur when two or more vehicles go into an infinite loop. An effective solution for avoiding deadlock is proposed to deal with this problem.

The main contributions of this research include: 1) The new approach dealing with the real-time problem with dynamic traffic flow. 2) The approach controls the passing at intersection of vehicles via wireless communication, so the system will be more feasible and flexible to implement in reality. 3) Our approach is more less complexity computation than others optimistic approaches. 4) The simulation shows that our results improve the waiting time of each vehicles at intersection.

2 Traffic Connected Vehicle Model and Notations

We assume a single intersection with four directions which leading to the road intersection are North, East, South and West which are marked as ($\mathcal{N}$), ($\mathcal{E}$), ($\mathcal{S}$) and ($\mathcal{W}$), respectively. Each direction has two lanes in the incoming direction, which are go-forward ($\mathcal{F}$) and turn-left ($\mathcal{L}$). Notice that in this paper, we do not consider for turn-right direction, since for almost kind of crossroads, the vehicle is able to take right-direction without any obstructions. By this way, a lane where the vehicle is passing at the intersection can be determined by a path ($\mathcal{P}$) of $\{ \mathcal{N}, \mathcal{E}, \mathcal{S}, \mathcal{W} \}$ and a direction $\mathcal{D}$ of $\{\mathcal{L}, \mathcal{F}\}$. As result, there is total eight lanes which operating in the model that are described by the pair ($\mathcal{P}$,$\mathcal{D}$): $\{ \mathcal{NL}, \mathcal{NF}, \mathcal{EL}, \mathcal{EF}, \mathcal{SL}, \mathcal{SF}, \mathcal{WL}, \mathcal{WD}\}$.

We can imagine the relationships of all lanes which include concurrent and conflicting lanes. In this regards, it is up to status of relationship among directions, vehicles will be allowed to operate (concurrency) or inoperative (conflict) to one another. According to previous traffic rules, to avoiding the conflict among vehicles at intersection, when the vehicles are passing, the system will lock all the conflicting lanes by controlling traffic light system (red light). However, in this case, the vehicles which in the different paths can not move even there are some lanes which are concurrent lanes. For instance, when a vehicle in lane $\mathcal{NF}$ is passing, all remaining three paths have to be blocked, although the lane $\mathcal{SF}$ in path $\mathcal{S}$ is concurrent with lane $\mathcal{NF}$. In order to deal with this problem, the existing systems allow the vehicles in the paths $\mathcal{N}$ and $\mathcal{S}$ pass together (the same with paths $\mathcal{W}$ and $\mathcal{E}$). Even though, the conflicts among vehicles sometimes still occur since the vehicles in conflicting lanes pass simultaneously (e.g., $\mathcal{NF}$ and $\mathcal{SL}$).

2.1 Connected Vehicle Architecture

In connected vehicle concept, the vehicles can be identified based on their ID (e.g., the plate number of the vehicle). The vehicle communicate with infrastructure (V2I) and other vehicles (V2V) under the ID as the address. Each vehicle is equipped the positioning system (VPS) to determine its location. In this regards, the traffic management system which located in the intersection is able to get the information of vehicles in term of their location and directions when they move into the intersection area. Fig. 1(a) shows the architecture of the traffic connected-vehicle at intersection. Sensors are developed in each path to determine the vehicles move into the simulation area which follow a certain random distribution. The wireless communication

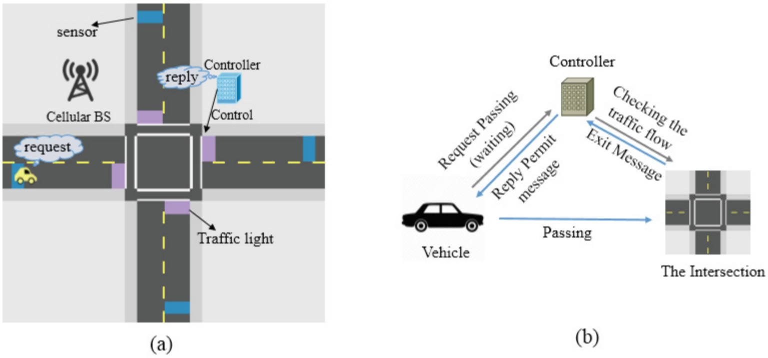


Fig. 1 (a) Traffic Connected Vehicle Architecture; (b) Interaction between Vehicle and Controller

devices (Cellular Base Station) are designed in the intersection to enable the communications between vehicle and others. We assume that the transmission range of wireless devices are able to cover all of the intersection area. Moreover, the wireless channel is FIFO channel, it means the traffic management system will receive the request from vehicles as FIFO model.

3 Process Synchronization Approach

As we mention above, the existing approaches for management traffic flow at intersection still have some problems, especially in case of high traffic volume such as traffic congestion, the large number of waiting vehicles in queue lanes, and vehicle crash at intersection because of conflict directions of vehicles. In this paper, we propose a new approach to deal with those problems based on recent advanced vehicle technologies. The idea of the proposed approach based on synchronization theory. In computer science, process synchronization refers to the idea that multiple processes are to join up at a certain point, in order to reach an agreement or commit to a certain sequence of action. This approach can use for controlling the traffic flow at intersection when multiple vehicles pass the intersection at the same time. Fig. 1(b) shows the interaction between vehicle and controller for passing the intersection of vehicle. By this way, when a vehicle want to pass the intersection, they have to send a request to the controller. Based on the traffic flow at intersection, the controller will reply the permit passing message to vehicle or put the request in the pending list, and the vehicle have to wait until receive reply from controller. After passing the intersection, the vehicle has to send the release message to Controller. The algorithms for this process are proposed in below.

3.1 Algorithms for Controlling Traffic Flow

The Algorithms 1 and 2 show the communication between vehicle and controller. When a vehicle i moves into the simulation area, it sends a request message with id and its lane li to controller and waiting until receiving the acknowledgement (reply message) from controller. On the other hand, when the controller receives the request message from vehicle i, it will check the traffic flow in the core area of the intersection. If in core area, there is a vehicle j which from conflicting lane with the lane of vehicle i is moving, or if the deadlock occurs when vehicle i pass into the intersection , then it will put the vehicle i in the pending list Pl and block the path of vehicle i based on traffic light (red-light). In other case, it sends a rely message to vehicle for permitting the passing.

Algorithm 1. Messages of Vehicles

```
Function Request-message
if vehicle i arrive intersection then
|  Send request-mesg (idi,li) ;
|  Sti := waiting;
end
Function Receive-message
while pas = False do
|  Sti:= Passing;
end
if pas = True then
|  Send release-measg (idi); Sti:= exiting;
end
```

When the vehicle already exits the intersection, it has to send a release message to the controller. After receiving the release message, the controller checks the pending list and then send the reply message. In case of more than one waiting reply message in the pending list, the controller will check one by one based on FIFO model.

Algorithm 2. Messages of Controller

```
Function Receive-Request (idi,li)
if no conflict and check-deadlock(i,li,G(N, E)) = False then
|  Send Reply-mesg(idi);
else
|  Block Pi; Add i to the pending list Pl;
end
Function Receive-Release(idi)
for each j of pending list Pl do
|  if no conflict and check-deadlock(j) = False) then
|  |  Send Reply-mesg(idj);
|  |  Remove j to the pending list Pl;
|  end
end
```

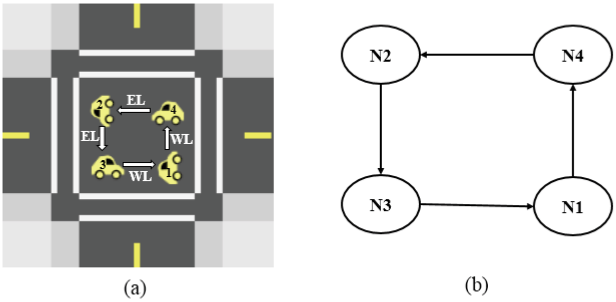


Fig. 2 (a) The Traffic Deadlock; (b) The Cycle Graph

3.2 Deadlock Avoidance Algorithm

Although we are able to manage the conflict among vehicles by using the process synchronization approach which only allows the vehicles in concurrent lanes passing the intersection simultaneously. However, in case of high traffic flow, the traffic deadlock sometimes still happens when two or more vehicles are each waiting for the other. Fig. 2(a) shows an example of the deadlock. The vehicles go into an infinite loop since they have to wait to each other. For almost existing systems, they assume that deadlock never occurs because the time intervals between occurrences of traffic deadlock are large and the time for waiting is tolerable. For connected vehicles technology, the traffic management system is able to collect all necessary information from vehicles for estimating whether the deadlock happens or not at the certain time when a vehicle pass into the intersection. By this way, the system can be detected the deadlock by using wait-for graph algorithm [15].

The process of the detection deadlock algorithm is given as follow: 1) first, we transform the deadlock detection problem into wait-for graph $G(N, E)$ problem. By this way, the vehicles are represented as nodes, an edge from vehicle i to vehicle j implies j is holding a place that i need and thus i waiting for j pass that place. 2) Then, using finding-cycle algorithm in a directed graph to detect the cycle in graph. 3) The deadlock will be detected if the result is True. In contrast, the vehicle can pass the intersection. Fig. 2(b) shows a wait-for graph with a cycle of the example in Fig. 2(a). The complexity computation of this algorithm is $\mathcal{O}(n^2)$ where n is total number of vehicle.

4 Performance Evaluation

To evaluate the effectiveness of the proposed approach, in this section, we take a comparison our system with the existing system using Round-Robin scheduling algorithm for controlling traffic light. Although there have been some proposed algorithms

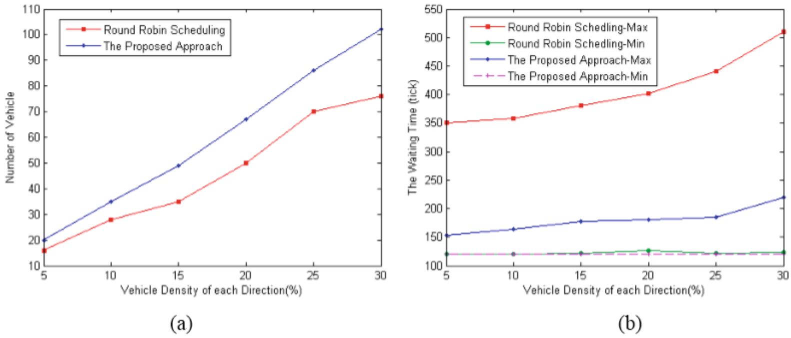
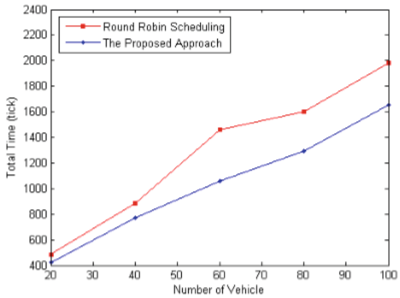


Fig. 3 (a) Number of vehicles pass the intersection at a certain time; (b) Maximum and Minimum waiting time of vehicle.

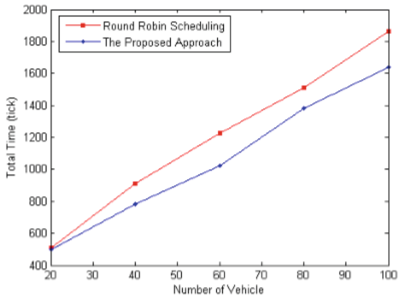
focus on dynamic traffic light system, however, as we mention in introduction section, since their computational complexity, it is infeasibility to real-time control. In contrast, our approach is able to implement based on connected technologies among objects. In this regards, we measure the performance of the traffic control algorithms in term of increasing density of vehicle which vehicles arrive to the intersection according to random distribution in each direction of the intersection.

Fig. 3(a) shows the number of vehicles can pass intersection at a certain time (1000 ticks). The number of vehicle density indicates percentage of vehicle arriving at intersection in each direction at a certain time (10 ticks). As we can see, our outperform is better than existing system since our system is a dynamic traffic light system in real-time, the vehicle can pass into the intersection anytime when there has not any vehicle in conflicting lanes is passing. Consequently, we can reduce the waiting time of each vehicle. For more details, fig. 3(b) shows the maximum and minimum waiting time of vehicle. The waiting time of vehicle obviously increase when the density of vehicle get higher. However, in our approach, the change does not vary too different. The reason is that in our system, the priority of request messages from vehicle is FIFO model, so the vehicle is able to pass into the intersection as soon as possible. The waiting time only get high when the traffic flow is overloaded.

In above experiment, we assume that the arrival rate of vehicles at intersection in all direction is similar. To evaluate more general of the proposed approach, the different arrival rate is also considered. Fig. 4, shows total time that all vehicles are able to exit the simulation area in term of same arrival rate of lanes (fig. 4(a)) and different arrival rate (fig. 4(b)) where the volume in west-east road is two times compare with north-south road. As a results, we can see our approach is better than the existing system in case of both uniform volume and non-uniform volume.



(a)



(b)

Fig. 4 (a) Total time for all vehicles pass the Intersection with same Arrival Rate; (b) Total time for all vehicles pass the Intersection with different Arrival Rate.

5 Conclusion and Future Works

In this paper, inspired of advanced vehicle technologies, we introduce a new approach for real-time management the traffic flow based on process synchronization approach. In this regards, when the vehicle passes into the intersection area, they can communicate with others to avoid the conflict between them, so that they are able to improve their waiting time in case of high traffic flow. Moreover, in this work, we also propose a new algorithm for avoiding the traffic deadlock which was not considered in existing systems. The results show that our approach can reduce the waiting time of vehicles when they pass into the intersection. The proposed system is feasible and flexible for development based on advanced vehicle technologies. However, to implement in reality where there have various shapes of intersections. In this regards, for our future works, we are interested to extend the model to the multiple crossroads e.g. multiple intersections, roundabout-intersection model and so on.

Acknowledgement This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (NRF-2014R1A2A2A05007154). Also, this research has been partially supported by Spanish Ministry of Economy and Competitivity and European Regional Development Fund FEDER (TIN2014-56494-C4-4-P), and by Comunidad Autonoma de Madrid under project CIBERDINE S2013/ICE-3095.

References

1. Directive 2010/40/EU of the European Parliament and of the Council of July 7, 2010
2. Chen, X., Shi, Z.: Real-coded genetic algorithm for signal timings optimization of a signal intersection. In: Proceedings of the First International Conference on Machine Learning and Cybernetics (2002)

3. Gokulan, B.P., Srinivasan, D.: Distributed Geometric Fuzzy Multi-agent Urban Traffic Signal Control. IEEE Transactions on Intelligent Transportation Systems **11**(3), 714–727 (2010)
4. Qiao, J., Yang, N.D., Gao, J.: Two-Stage Fuzzy Logic Controller for Signalized Intersection. IEEE Transactions Systems, Man and Cybernetics- Part A: Systems and Humans **41**(1), 178–184 (2011)
5. Srinivasan, D., Choy, M.C., Cheu, R.L.: Neural Networks for Real-Time Traffic Control System. IEEE Transactions Intelligent Transportation Systems **7**(3), 261–272 (2006)
6. Zhao, D., Dai, Y., Zhang, Z.: Computational Intelligence in Urban Traffic Signal Control: A Survey. IEEE Transactions Systems, Man, and Cybernetics-Part C: Applications and Reviews **42**(4), 485–494 (2012)
7. Zhao, L., Peng, X., Li, L., Li, Z.: A Fast Signal Timing Algorithm for Individual Oversaturated Intersections. IEEE Transactions Intelligent Transportation Systems **12**(1), 280–283 (2011)
8. Wu, W., Zhang, J., Lua, A., Cao, J.: Distributed Mutual Exclusion Algorithms for Intersection Traffic Control. IEEE Transactions On Parallel and Distributed Systems **26**(1), 65–74 (2015)
9. Lai, Y., Zheng, Y., Cao, J.: Protocols for traffic safety using wireless sensor network. In: Lecture Notes in Computer Science, vol. 4494, pp. 37–48 (2007)
10. Iqbal, Z.: Self-organizing wireless sensor networks for inter-vehicle communication. Master Thesis, Department of Computer and Electrical Engineering, Halmstad University (2006)
11. Khalil, M.Y., Jamal, N.A., Ali, M.S.: Intelligent Traffic Light Flow Control System Using Wireless Sensors Networks. Journal of Information Science and Engineering **26**, 753–768 (2010)
12. Zhou, B., Cao, J., Zeng, X., Wu, H.: Adaptive traffic light control in wireless sensor network-based intelligent transportation system. In: Proceedings of the IEEE Vehicular Technology Conference Fall, September 2010
13. Bento, L.C., Parafita, R., Santos, S., Nunes, U.: Intelligent traffic management at intersections: legacy mode for vehicles not equipped with V2V and V2I communications. In: Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation System (ITSC 2013), Netherlands (2013)
14. Lu, N., Cheng, N., Zhang, N., Shen, X., Mark, J.W.: Connected Vehicles: Solutions and Challenges. IEEE Internet of Things Journal **1**(4), 289–299 (2014)
15. Silberschatz, A., Galvin, P., Gagne, G.: Operating System Concepts. John Wiley and Sons, INC., p. 260. ISBN 0-471-25060-0 (2003)

Developing an Individualized Survival Prediction Model for Colon Cancer

Ana Silva, Tiago Oliveira, Paulo Novais, José Neves and Pedro Leão

Abstract In this work a 5-year survival prediction model was developed for colon cancer using machine learning methods. The model was based on the SEER dataset which, after preprocessing, consisted of 38,592 records of colon cancer patients. Survival prediction models for colon cancer are not widely and easily available. Results showed that the performance of the model using fewer features is close to that of the model using a larger set of features recommended by an expert physician, which indicates that the first may be a good compromise between usability and performance. The purpose of such a model is to be used in Ambient Assisted Living applications, providing decision support to health care professionals.

1 Introduction

Colorectal cancer or bowel cancer is a pathology that affects the lower portion of the gastrointestinal tract. This is the most common cancer of the digestive system and the fourth most lethal with a mortality rate of 6.41% [7]. Although colon and rectal cancers are considered to be very similar pathologies, the truth is they appear in anatomically different regions, they may be associated with different genetic causes, and may progress differently according to distinct molecular pathways and interactions, thus requiring different treatments [16]. The work disclosed herein focuses solely on colon cancer, which may develop in the cecum, ascending colon, transverse colon, descending colon, and sigmoid. Surgical resection is the primary treatment

A. Silva · T. Oliveira(✉) · P. Novais · J. Neves
Algoritmi Centre/Department of Informatics, University of Minho, Braga, Portugal
e-mail: a55865@alunos.uminho.pt, {toliveira,pjon,jneves}@di.uminho.pt

P. Leão
ICVS/3B's - PT Government Associate Laboratory, Braga/guimarães, Portugal
e-mail: pedroleao@ecsaude.uminho.pt

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_10

for early stages of colon cancer. The accurate prediction of survival is important for patients with cancer so that they can make the most out of the rest of their lives. It is also important to help clinicians to make the best decisions for patients and it is essential for palliative care. Therefore, the objectives of this work are the following: i) to make an individualized prediction of the survivability of a colon cancer patient in each year of the five years following treatment; ii) to determine the ideal number of features necessary for an accurate prediction; and iii) to determine which features are the most important for survival prediction of colon cancer patients. The number of features is important in order to make the prediction model available in a clinical decision support application, which is the end goal of the work. If a physician has to provide too many inputs, thus making the task of using the application difficult and time-consuming, he may lose interest and not use the tool at all. Such a tool would assist a health care professional in his daily practice.The prediction model was developed using data from the Surveillance, Epidemiology, and End Results (SEER) program [12], a large cancer registry in the United States, and arguably the most complete cancer database in the world.

The paper is structured as follows. Section 2 describes related work in colon cancer survival prediction. Section 3 describes the steps and machine learning methods used to develop the prediction model. The corresponding experimental results are disclosed and discussed in Section 4. Finally, Section 5 provides concluding remarks about the work done so far and future work considerations.

2 Related Work

Most of the existing approaches for colon cancer survival prediction are based on the SEER data. An example is the web-based calculator developed in [4] whose underlying prediction model is the Nodes + Prognostic Factors (NAP), based on the number of positive lymphatic nodes combined with other prognostic features. The model has an underlying biological motivation, reflected in the use of the probability of a cancerous cell invading healthy tissues to formulate equations for cancer lethality, combined with other prognostic features estimated by means of simulation of several statistical tests. The model requires inputs for 9 features and provides a prediction of the mortality risk over the period of 15 years. Another SEER-based approach is the one followed in [5], also made available in the form of a web application. The prediction model has 5 input features, derived through a Cox regression analysis to evaluate simultaneous effects of multiple variables on survival. This resulted in adjusted survival functions stratified by 5 features. The conditional survival probabilities for a period of 10 years produced by the model are calculated on the basis of the adjusted survival functions for the features, controlled for the influence of other covariates in the final model. Another example, but with a strong machine learning component, is the work in [1], in which a 5-year survival prediction model was developed using ensemble machine learning with supervised classification. A total of

13 features were selected in that work, and the resulting model achieved an overall high performance in terms of accuracy and receiver operating characteristic (ROC).

The work developed herein distances itself from the works in [4, 5] by treating survival prediction as a classification problem and applying varied machine learning methods to obtain a model capable of individualized survival prediction. In this regard, it is influenced by the methodology followed in [1], whose work will serve as a reference for direct comparison. At the same time, this work aims to produce predictions using fewer features than the existing approaches.

3 Development of the Prediction Model

The survival prediction system for colon cancer should be able to accept a number of inputs for selected prediction features and, for each of the 5 years following treatment, produce an output stating whether the patient in question will survive that year or not, along with a confidence value for the prediction. The development of a prediction model capable of this required several phases, from the preprocessing of SEER data to the selection of the best model. The workflow that leads to the learning and evaluation of the model comprises the following phases by order of occurrence: Preprocessing, Split Dataset, Balancing Data, Feature Selection, Modeling, and Evaluation. The software chosen to develop the prediction model was RapidMiner, an open source data mining soft- ware. It has a workflow-based interface that allows an intelligible construction of complex data management processes. Moreover, it offers an intuitive application programming interface (API). It is important to clarify that, given that survival prediction was handled as a classification problem, five classification models for each year were developed. Theses models were posteriorly combined, in a programmatic manner, into a model capable of providing a prediction for each year with a single interaction.

3.1 Preprocessing, Split Dataset, and Balancing Data

The colorectal cancer data from SEER contained 515,791 records and consisted of 146 attributes, some of them only applicable to a limited period within the time of data collection. The data was reduced to 38,592 records after the preprocessing phase and selecting the colon cancer patients.

During the Preprocessing phase, it was defined that the period of interest would be from 2004 onwards in order to minimize the occurrence of missing data due to the applicability of the attributes. Additionally, empty attributes, attributes that are not applicable to this type of cancer, and attributes that are not directly related with the vital status of the patient were removed (e.g. the number identifying the registry of the patient). Only the adult patients (with age greater than or equal to 18 years old) were selected for further processing and the missing values were replaced with

Table 1 Class distribution for each target label in the sub-datasets.

	Target Labels				
	1 Year	2 Year	3 Year	4 Year	5 Year
Not Survived	24.51%	32.60%	36.96%	39.35%	41.07%
Survived	75.49%	67.40%	63.04%	60.65%	58.93%

the *unknown* code. Patients who had a survival time inferior to 60 months (5 years), the maximum period for which the model under development is supposed to predict survival, and those who passed away of causes other than colon cancer were sampled out from the training set, as their inclusion was considered to be of little relevance to the problem. The numeric attributes were converted to nominal and the binary classes (*survived* and *not survived*) were derived for the target labels 1-, 2-, 3-, 4- and 5-year survival. Based on existing attributes, new ones, such as the number of regional lymph negative nodes, the ratio of positive nodes over the total examined nodes and also patient relapse, were calculated. After the Preprocessing, the attributes were reduced to 61, including the new attributes and the target labels.

In the Split Dataset phase, the data was divided into five sub-datasets, split by target label, according to the corresponding survival year. Table 1 shows the class distribution in each sub-dataset. As observed, the classes are not equally represented. The problem of using imbalanced datasets is important, from both the algorithmic and performance perspectives. An overview of classification algorithms for the resolution of this kind of problem [9] concluded that hybrid sampling techniques, i.e., combining over-sampling of the minority class with under-sampling of the majority class, can achieve better performances than just oversampling or undersampling. As such, in the Balancing Data phase, hybrid sampling, as described in [9], was applied in order to generate balanced sub-datasets with 38,592 records each.

3.2 Feature Selection

The Feature Selection phase was crucial to determine the most influential features on the survival of colon cancer patients. In order to accomplish this the Optimize Selection operator [14] of RapidMiner was used. It implements a deterministic and optimized selection process with decision trees and *forward selection*. The process was applied to each sub-dataset for the target label. Only the features in common to all the sub-datasets were selected and used to construct the prediction models. Table 2 shows the selected features and their meaning. The 6 selected features were compared with a set of 18 features (shown in Table 3) indicated by a specialist physician on colorectal cancer. These two sets of features were mapped to attributes in the sub-datasets and later used to generate and evaluate the prediction models.

Table 2 Attributes selected in the Feature Selection process.

Attribute	Description
Age recode with < 1 year old	Age groupings based on age at diagnosis (single-year ages) of patients (< 1 year, 1-4 years, 5-9 years, ..., 85+ years)
CS Site-Specific Factor 1	Interpretation of the Carcinoembryonic Antigen (CEA) test results
CS Site-Specific Factor 2	The clinical assessment of regional lymph nodes
Derived AJCC Stage Group	The grouping of the TNM information combined
Primary Site	Identification of the site in which the primary tumor originated
Regional Nodes Examined	The total number of regional lymph nodes that were removed and examined by the pathologist

Table 3 Attributes selected by a specialist physician on colorectal cancer.

Attribute	Description
Age at Diagnosis	The age of the patient at diagnosis
CS Extension	Extension of the tumor
CS Site-Specific Factor 8	The perineural Invasion
CS Tumor Size	The size of the tumor
Derived AJCC T, N and M	The AJCC T, N and M stage (6th ed.)
Grade	Grading and differentiation codes
Histologic Type	The microscopic composition of cells and/or tissue for a specific primary
Laterality	The side of a paired organ or side of the body on which the reportable tumor originated
Primary Site	*
Race Recode (White, Black, Other)	Race recode based on the race variables
Regional Nodes Examined	*
Regional Nodes Positive	The exact number of regional lymph nodes examined by the pathologist that were found to contain metastases
Regional Nodes Negative	(Regional nodes examined - Regional nodes positive)
Regional Nodes Ratio	(Regional nodes negative over Regional nodes examined)
Relapse	The relapse of the patients for colon cancer
Sex	The sex of the patient at diagnosis

* Described in Table 2.

3.3 *Modeling and Evaluation*

The classification strategies used in the Modeling phase consisted mostly of ensemble methods. The classification schemes applied were meta-classifiers. This type of classifier is used to boost basic classifiers and improve their performance. All the possible combinations of the classifiers were explored, according to the algorithms and type of attributes allowed. The tested meta-classifiers were:

– **Bagging** [3]: Also called bootstrap aggregating. It splits the data into *m* different training sets on which *m* classifiers are trained. The final prediction results from the equal voting of each generated model on the correct result.

- **AdaBoost** [8]: This meta-classifier calls a new weak classifier at each iteration. A weight distribution which indicates the weight of examples in the classification is updated. It focuses on the examples that have been misclassified so far in order to adjust subsequent classifiers and reduce relative error.
- **Bayesian Boosting** [14]: A new classification model is produced at each iteration and the training set is reweighed so that previously discovered patterns are sampled out. The inner classifier is sequentially applied and the resulting models are later combined into a single model.
- **Stacking** [6]: This meta-classifier is used to combine base classifiers of different types. Each base classifier generates a model using the training set, then a meta-learner integrates the independently learned base classifier models into a high level classifier by re-learning a meta-level training set.
- **Voting** [10]: Each inner classifier receives the training set and generates a classification model. The prediction of an unknown example results from the majority voting of the derived classification models.

A group of basic classifiers were selected to be used in ensembles with the above-described meta-classifiers. The group includes some of the most widely used learners [14] available in RapidMiner, namely the k-NN, the Naive Bayes, the Decision Tree, and the Random Forest. A total of fourteen classification schemes were explored for each set of attributes (6 and 18 attributes) for 1, 2, 3, 4, and 5 survival years. The learning combinations of meta-classifiers with basic classifiers are as follows. The Stacking model used k-NN, Decision Tree, and Random Forest classifiers as base learners, and a Naive Bayes classifier as a Stacking model learner. The Voting model used k-NN, Decision Tree and Random Forest as base learners. The other models were used in combination with each basic classifier. For evaluation purposes, 10-fold cross-validation [15] was used to assess the prediction performance of the generated prediction models and avoid overfitting.

4 Experimental Results and Discussion

Each classification scheme was evaluated using the prediction accuracy and the area under the ROC curve (AUC) for 1, 2, 3, 4, and 5 years. The accuracy is the percentage of correct responses among the examined cases [2]. The AUC is the percentage of randomly drawn data pairs of individuals that have been accurately classified in the two populations [11]. Tables 4 and 5 present all the results obtained for prediction accuracy and AUC respectively of the top three algorithms. The average performances in terms of accuracy and AUC of the top three learning schemes for the 5 years are shown in Figures 1a and 1b respectively. From the observation of the figures and the tables, it is obvious that almost all the classification methods demonstrated high performances, particularly the ones using decision trees. Out of those, the Stacking models showed a slightly better average performance both in terms of accuracy (Figure 1a) and AUC (Figure 1b).

Table 4 Survivability Percentage Accuracy.

	Accuracy											
	1 Year		2 Year		3 Year		4 Year		5 Year		Average	
Ensemble Model	18 attributes	6 attributes	18 attributes	6 attributes	18 attributes	6 attributes	18 attributes	6 attributes	18 attributes	6 attributes	18 attributes	6 attributes
Stacking	98.28%	96.15%	97.63%	96.78%	98.02%	97.12%	98.02%	97.26%	97.83%	96.81%	97.96%	96.82%
Bayesian Boosting with Decision Tree	97.83%	96.33%	97.53%	96.76%	97.81%	96.95%	97.84%	96.98%	97.85%	96.72%	97.77%	96.75%
AdaBoost with Decision Tree	97.83%	96.35%	96.89%	96.78%	97.81%	96.95%	97.84%	97.02%	97.85%	96.74%	97.64%	96.77%

Table 5 Survivability AUC.

	AUC											
	1 Year		2 Year		3 Year		4 Year		5 Year		Average	
Ensemble Model	18 attributes	6 attributes	18 attributes	6 attributes	18 attributes	6 attributes	18 attributes	6 attributes	18 attributes	6 attributes	18 attributes	6 attributes
Stacking	0.991	0.984	0.993	0.987	0.994	0.99	0.994	0.991	0.994	0.991	0.993	0.989
Voting	0.988	0.979	0.988	0.982	0.989	0.983	0.99	0.985	0.988	0.984	0.989	0.983
Bayesian Boosting with Decision Tree	0.977	0.963	0.984	0.97	0.979	0.969	0.984	0.973	0.986	0.967	0.982	0.9684

(a) Average survivability percentage accuracy. (b) Average survivability percentage AUC.

Fig. 1 Comparison of the 18-attribute models with the 6-attribute models.

Comparing the results of the 6-attribute stacking models with those of the 18-attribute models, it is possible to say that the differences are small. With an average of 96.82% for accuracy and 0.989 for AUC, the 6-attribute stacking models had prediction accuracies for years 1 to 5 of 96.15%, 96.78%, 97.12%, 97.26% and 96.81% (as seen in Table 4), and AUCs of 0.984, 0.987, 0.990, 0.991 and 0.991 (as seen in Table 5). The 18-attribute models had an average accuracy of 97.96%, with values for years 1 to 5 of 98.28%, 97.63%, 98.02%, 98.02% and 97.83%. The average AUC was 0.993, and the remaining values were 0.991, 0.993, 0.994, 0.994 and 0.994, for years 1 to 5. It should be noted that, in addition to the close performances, the difference between the number of attributes used is important. The results show that it is possible to build a model with less than half of the features indicated by the expert physician. Regarding the attributes obtained in the feature selection process, with the exception of the site-specific factors, they were all connected with the features indicated by the specialist physician. Comparing this approach with others mentioned in Section 2, fewer features were necessary to develop the prediction model. Moreover, in the approach followed in [1], the closest to the one followed herein, the best model of colon cancer survival prediction was based on a Voting classification scheme, with prediction accuracies of 90.38%, 88.01%, and 85.13% and AUCs of 0.96, 0.95, and 0.92 for years 1, 2 and 5. As such, the present work represents an improvement and was able to achieve better results.

5 Conclusions and Future Work

The best meta-classification scheme is a Stacking model combining k-NN, Decision Tree, and Random Forest classifiers as base learners and a Naive Bayes classifier as a stacking model learner. The ideal number of features for colon cancer survival prediction was found to be 6. The selected set includes: age, CS site-specific factor 1, CS site-specific factor 2, derived AJCC stage group, primary site, and regional nodes examined. Overall the developed model was able to present a good performance with fewer features than most of the existing approaches.

Since rectal cancer shares a lot in common with colon cancer, a similar analysis is underway. Additionally, a mobile application to make the model available to the health care community is under development. One intends to have this clinical decision support application available in different platforms, ready to assist health care professionals in carrying out their duties at any time. In order to ensure that the model is able to adapt and adjust, an on-line learning scheme is also being prepared. It will be possible for users to dynamically feed new cases to the prediction system and make it change in order to provide better survival predictions. This type of model could also prove to be very useful when integrated in computer-interpretable guideline systems, such as the one described in [13], as a way to provide dynamic knowledge to rule-based decision support.

Acknowledgements This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2013. The work of Tiago Oliveira is supported by a FCT grant with the reference SFRH/BD/85291/2012.

References

1. Al-Bahrani, R., Agrawal, A., Choudhary, A.: Colon cancer survival prediction using ensemble data mining on seer data. In: 2013 IEEE International Conference on Big Data, pp. 9–16 (2013)
2. Bradley, A.P.: The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition **30**(7), 1145–1159 (1997)
3. Breiman, L.: Bagging predictors. Machine Learning **24**(2), 123–140 (1996)
4. Bush, D.M., Michaelson, J.S.: Derivation: Nodes + PrognosticFactors Equation for Colon Cancer. Tech. rep. (2009)
5. Chang, G.J., Hu, C.Y., et al.: Practical application of a calculator for conditional survival in colon cancer. Journal of Clinical Oncology **27**(35), 5938–5943 (2009)

6. Džeroski, S., Ženko, B.: Is combining classifiers with stacking better than selecting the best one? Machine Learning **54**(3), 255–273 (2004)
7. Ferlay, J., Soerjomataram, I., Ervik, M., et al.: Globocan 2012: Estimated cancer incidence, mortality and prevalence worldwide in 2012 (2012). http://globocan.iarc.fr (last visited on December 27, 2015)
8. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. **55**(1), 119–139 (1997)
9. Ganganwar, V.: An overview of classification algorithms for imbalanced datasets. Int. J. Emerg. Technol. Adv. Eng **2**(4), 42–47 (2012)
10. Kittler, J.: Combining classifiers: A theoretical framework. Pattern Analysis and Applications **1**(1), 18–27 (1998)
11. Klepac, G., Klepac, G., et al.: Developing Churn Models Using Data Mining Techniques and Social Network Analysis, 1st edn. IGI Global, Hershey (2014)
12. National Cancer Institute: Surveillance, epidemiology and end results program (2015). http://seer.cancer.gov/data/ (last visited on October 01, 2015)
13. Oliveira, T., Leão, P., Novais, P., Neves, J.: Webifying the computerized execution of clinical practice guidelines. In: Bajo Perez, J., Corchado, J.M., et al. (eds.) The PAAMS Collection SE - 18, Advances in Intelligent Systems and Computing, vol. 293, pp. 149–156. Springer International Publishing (2014)
14. RapidMiner: Rapidminer documentation: Operator reference guide (2016). http://docs.rapidminer.com/studio/operators/ (last visited on March 01, 2016)
15. Refaeilzadeh, P., Tang, L., Liu, H.: Cross-validation. In: Liu, L., Özsu, M. (eds.) Encyclopedia of Database Systems, pp. 532–538. Springer US (2009)
16. Yamauchi, M., Lochhead, P., Morikawa, T., et al.: Colorectal cancer: a tale of two sides or a continuum? Gut **61**(6), 794–797 (2012)

Providing Advanced Touristic Services by Means of Augmented Reality and Multimodal Dialog

David Griol and José Manuel Molina

Abstract In this paper, we describe a mobile application that combines Augmented Reality and Multimodal Dialog Systems to provide advanced touristic services. We have coupled machine vision solutions and sensors currently available in mobile devices to provide Augmented Reality functionalities. This integration facilitates these devices perceiving the surrounding world in similar ways as humans and providing context-aware information and services based on the location of the user. In addition, Multimodal Dialog Systems improve the access to these services by making them more accessible through a more intuitive, natural, and transparent human-machine interface. The paper also presents the results of a preliminary assessment of the developed App that shows the benefits of our proposal.

Keywords Human-Computer Interaction · Augmented Reality · Multimodal Interfaces · Spoken interaction · Mobile applications

1 Introduction

Recent advances of mobile technology are increasingly activating Ubiquitous Computing and Ambient Intelligence, which involve important objectives like technology services being available anywhere and anytime, minimal user distraction, and integration of the strengths of both the real and digital worlds in user single interfaces, which are generally efficient, effortless and intuitive for people [3, 5, 7, 10].

The achievement of these objectives are intimately related to the number of machine vision solutions and sensors currently integrated in mobile devices (e.g., cameras, microphones, and orientation sensors), which allow these devices to perceive

D. Griol(✉) · J.M. Molina
Computer Science Department, Carlos III University of Madrid,
Avda. de la Universidad, 30, 28911 Leganés, Spain
e-mail: {david.griol,josemanuel.molina}@uc3m.es

© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_11

the surrounding world in similar ways as humans and provide services and information contents based on the location of the user. The improved hardware properties of mobile devices are also related to the advances in technologies connected with the described objectives, such as Augmented Reality and Multimodal Dialog Systems.

Augmented Reality (AR) complements the real world by means of dynamically overlaying pieces of context sensitive virtual information in a real setting, real time, and offering a 3-dimensional alignment between real and virtual objects [1, 4]. Although initially the application of this technology required high-end electronics hardware and sophisticated equipments, mobile devices with improved hardware properties have increased the uses of this technology originally restricted to laboratory environments. Handheld Augmented Reality (HAR) has thus the potential to introduce Augmented Reality for purposes that are common in mobile contexts and for large audiences due to the widespread use of suitable handheld devices [9].

Different technologies have recently emerged to facilitate the accessibility of handheld devices, which reduced size makes them difficult to operate in some situations and specially for some user groups. For example, Multimodal Dialog Systems [6, 8] can be employed to build more natural interaction with mobile devices by means of the combination of speech and visual information. They can be defined as computer programs designed to emulate communication capabilities of a human being including several communication modalities, such as speech, tactile and visual interaction.

In addition, these systems typically employ several output modalities to interact with the user, which allows to stimulate several of his senses simultaneously, and thus enhance the understanding of the messages generated by the system. For this reason, multimodal conversational agents are becoming a strong alternative to traditional graphical interfaces which might not be appropriate for all users and/or applications [8].

In this paper we describe an App for Android-based mobiles devices that combines the benefits of Mobile Augmented Reality and Multimodal Dialog Systems to provide advanced touristic services. The developed App employs an enhanced display paradigm stemming from the synergy of new mobile devices, context-awareness and AR, which has been enhanced by means of an effective and usable design based on a multimodal dialog system that combines spoken interaction with visual information for both input and output. The *Augmenta* App uses the camera and location services of mobile devices to locate and show touristic points of interest (POIs) around the user by means of Augmented Reality. Users can then select a specific POI, obtain information about it by means of spoken and visual inputs and outputs, access the wikipedia website to look up additional information, detect these points by means of image recognition, locate the nearest points and obtain guiding information to them from the current location.

2 The *Augmenta* App

Figure 1 shows the modular architecture of the *Augmenta* mobile application. Users can interact with the App by means of textual and tactile inputs (provided using physical or virtual keyboards and the screen) or by means of spoken interaction.

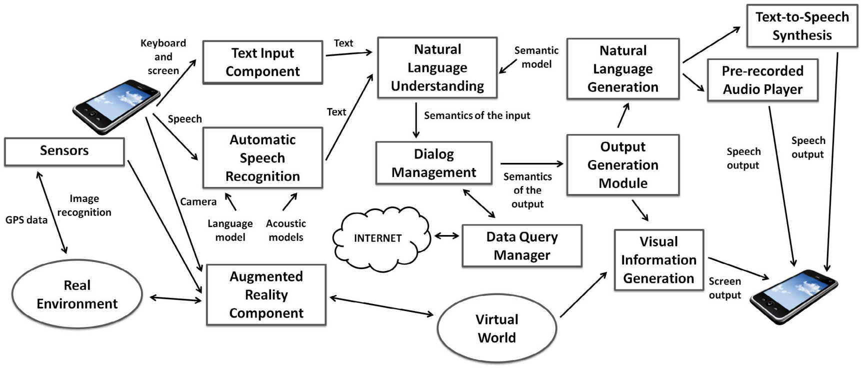


Fig. 1 Modular architecture of the *Augmenta* App

The developed App integrates the Google Speech API to include the speech recognition functionality. This functionality is available by means of a microphone icon on the Android keyboard, which activates the Google speech recognition service.Once the App has recognized what the user uttered, it is necessary to understand what he said. Natural language processing generally involves morphological, lexical, syntactical, semantic, discourse and pragmatical knowledge. A set of of grammars have been designed to carry the semantic interpretation of the user inputs.

The dialog manager of the system has to deal with different sources of information such as the NLU results, database queries results, application domain knowledge, and knowledge about the users and the previous dialog history to select the next system action. A statistical methodology that combines multimodal fusion and dialog management functionalities has been used to develop the dialog manager of the App [2]. The modality fission module receives abstract, modality independent presentation goals from the dialog manager. This module applies presentation strategies that decompose the complex presentation goal into presentation tasks. It also decides whether an object description is to be uttered verbally or graphically. The result is a presentation script that is passed to the the Visual Information and Natural Language generation modules.

The visual generation module creates the visual arrangement of the content using dynamically created and filled graphical layout elements. Since many objects can be shown at the same time on the display, the manager re-arranges the objects on the screen and removes objects, if necessary. The visual structure of the user interface (UI) is defined in an Android-based multimodal application by means of layouts.

Layouts can be defined by declaring UI elements in XML or instantiating layouts elements at runtime. Both alternatives can be combined in order to declare the application's default layouts in XML and add code that would modify the state of the screen objects at run time. Declaring the UI allows to better separate the presentation of the application from the code that controls its behavior.

As previously described, the essence of Augmented Reality is to display virtual contents in real images. The front webcamera of the mobile device is usually used to capture the true picture by the front webcamera and the location of the virtual contents must match the actual content as closely as possible. Advanced sensors and technologies are required to accurately display Augmented Reality. There are currently a number of AR SDK software solutions that can be used to develop apps for mobile devices, such as Vuforia, ARLab, BeyondAR, Catchoom, D'Fusion Mobile, Wikitude, SLARToolkit, etc.[1]

Although all of them offer advanced functionalities to develop AR mobile applications, we have selected the cross-platform Wikitude[2] SDK given that the contents are generated using very well-known web standards such as HTML 5, JavaScript and CSS; it supports working with geo-referenced data, image recognition and tracking functionalities; and provides a number of tutorials and sample projects fully documented.

The integration of the Wikitude SDK into the App is carried out by means of the ARchitectView class, which encapsulates the camera, graphics and web engines. The Augmented Reality parts of the application, which are called ARchitect World, are implemented by means of HTML pages that are overlapped to real images captured by the camera. The JavaScript language is used to access the different methods to detect geolocated places, incorporate images or videos, show distances, etc. Image recognition is based in Wikitude on the Target Management tool, which allows to create a target collection of PNG and JPEG images that can be used to detect and track particular places.

The Android sensor framework (*android.hardware* package) allows the SDK to access built-in sensors and acquire raw sensor data. Some of these sensors are hardware-based and some are software-based. Hardware-based derive their data by directly measuring specific environmental properties, such as acceleration, geomagnetic field strength, or angular change. Software-based sensors derive their data from one or more of the hardware-based sensors (e.g., linear acceleration and gravity sensors).

Android allows applications to access location services using the classes in the *android.location* package. The central component of the location framework is the *LocationManager* system service, also the Google Maps Android API permits to add maps to the application, which are based on Google Maps data. This API automatically handles access to Google Maps servers, data downloading, map display, and touch gestures on the map. The API can also be used to add markers, polygons and

[1] A detailed list compiling AR SDKs can be found at http://augmera.com/?p=461

[2] http://www.wikitude.com/developer

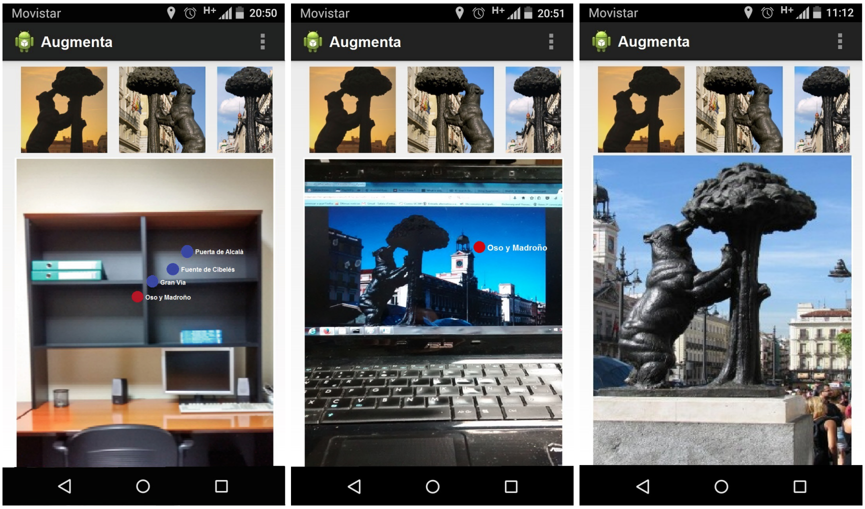


Fig. 2 Set of functionalities provided by the *Augmenta* App (I)

overlays, and to change the user's view of a particular map area. To integrate this API into an application, is it required to install the Google Play services libraries.

Figures 2 and 3 shows different screenshots of the *Augmenta* App with the different functionalities that it provides. Figure 2-left shows how the POIs around the user are displayed over the image that is captured by the camera of the mobile device. Users can select each one of them by means spoken or visual interaction with the App. Once a POI is selected, users can visualize photos of the corresponding location at the top of the screen. A database of the App includes the identifier, name, latitude, longitude, altitude, country, city, description, category, URL of the Wikipedia page, and collection of images of each one of the POIs for which the App can provide information.

Users can also listen to a summary explanation about the selected POI and ask questions related to this explanation, which are interpreted by means of the grammars designed for the NLU module of the application. Text-to-speech synthesis (TTS) is used to generate the voice signal used to transmit this explanation to the user. the Google TTS API to include the TTS functionality in an application. The *android.speech.tts* package includes the classes and interfaces required to integrate text-to-speech synthesis in an Android application. They allow the initialization of the TTS engine, a callback to return speech data synthesized by a TTS engine, and control the events related to completing and starting the synthesis of an utterance, among other functionalities.

Figure 2 also shows how the POI can also be automatically detected by means of the image recognition functionalities provided by the Wikitude SDK. The Wikitude Target Manager receives an image as input, which can be captured by the camera with a picture of the real place (Figure 2-center) or directly by means of using the

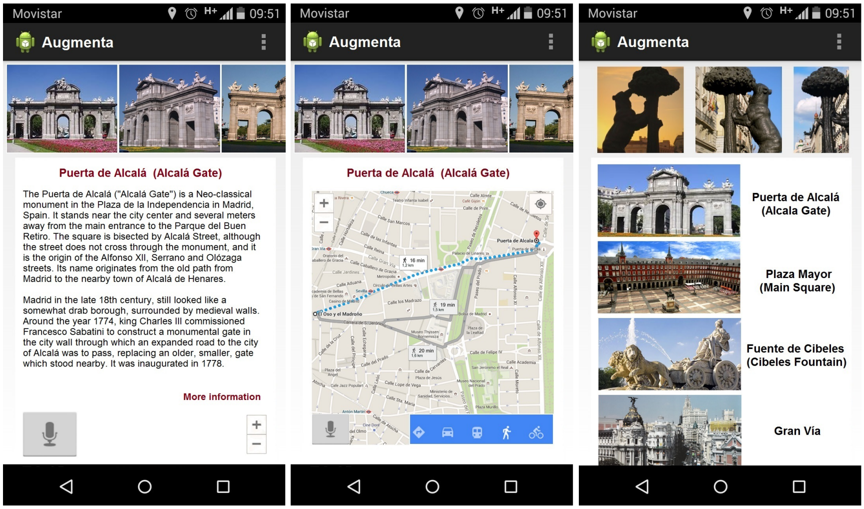


Fig. 3 Set of functionalities provided by the *Augmenta* App (II)

camera in the real place (Figure 2-right). A WTC file is generated as output, which is used by the JavaScript API of the Wikitude SDK to compare the captured image with the ones stored in the target collection for each POI.

Users can visualize a more detailed information about the selected POI by means of a screen in which a browser with the search results for the selected POI are shown or visualizing the information provided by the Wikipedia webpage for the selected object (Figure 3-left). Users can also access a screen showing the POIs that are near of the one that has been selected (Figure 3-center) and select them by means of visual or spoken interaction. They can also be guided to the different POIS. As Figure 3-right shows, the App displays a map centered in the current location of the user. The user can select the destination in three ways: spotting it with a finger in the map, introducing the address in a text field, or indicating it orally. In any of the three cases, the user is indicating a destination, which will be marked in the map by the system with a red sign. The system also shows the route to the destination from the current location. It can be shown visually or orally, and it is possible to set the preferred transportation means.

3 Preliminary Evaluation

A preliminary evaluation of the *Augmenta* App has been already completed with the participation of 30 recruited users aged from 17 to 45 (average age 27.3). The questionnaire with the following questions was defined for the evaluation: *Q1*: How much experience using smartphones do you have?; *Q2*: How much experience with

Table 1 Results of the evaluation of the *Augmenta* App by recruited users

	Min / max	Average	Std. deviation
Q1	4/5	4.37	0.59
Q2	2/5	3.11	1.47
Q3	4/5	4.23	0.65
Q4	4/5	4.85	0.11
Q5	3/5	4.07	1.28
Q6	4/5	4.83	0.37
Q7	4/5	4.89	0.17
Q8	4/5	4.03	0.66
Q9	4/5	4.53	0.41
Q10	4/5	4.81	0.35

applications with Augmented Reality do you have?; *Q3*: How understandable the steps required in the different functionalities of the App are?; *Q4*: How accurate is the information provided by the Augmented Reality functionalities of the App?; *Q5*: How well did the system understand you?; *Q6*: How well did you understand the system messages?; *Q7*: Was it easy for you to get the requested information?; *Q8*: If there were system errors, was it easy for you to correct them?; *Q9*: Was the interaction with the system quick enough?; *Q10*: In general, are you satisfied with the performance of the system?. The possible answers for each one of the questions were the same: *Never/Not at all*, *Seldom/In some measure*, *Sometimes/Acceptably*, *Usually/Well*, and *Always/Very Well*. All the answers were assigned a numeric value between one and five (in the same order as they appear in the questionnaire). The results of the questionnaire are summarized in Table 1.

As can be observed from the responses to the questionnaire, most of the users participating in the survey use smartphones and the Android OS. Few of them had a previous knowledge about Augmented Reality applications. Users positively evaluate the accuracy of the Augmented Reality functionalities as one of the most important benefits of the application. Most of the users found the application easy to use. The satisfaction with technical aspects of the application was also high, as well as the perceived potential to recommend its use. They also value the possibility of spoken interaction with the App, obtaining satisfactory results related to the understanding of the spoken information provided by the App. They also mentioned that the ambient noise could be the reason behind some misunderstandings of the multimodal interface. The global rate for the system was 8.4 (in the scale from 0 to 10).

4 Conclusions and Future Work

In this paper, we have described an App combining Augmented Reality functionalities and multimodal interaction by means of conversational interfaces. This combination provides a number of benefits for the development of mobile tourism applications: provide context-aware access to location-based information relevant to the immediate surroundings of tourists, enable access to variable content that can be timely and updated, deliver multimodal information sources including texts, videos, speech or images, and provide interactive annotations which can be integrated with map-based services and additional information.

In addition, the possibility of multimodal communication offers the user combinations of input and output modalities for interacting with the App, taking advantage of the naturalness of speech. In particular, this is a useful alternative to graphic user interfaces for mobile devices, allowing the use of other communication as an alternative to tapping through different menus.

As future work, we want to extend the developed App to answer more complex questions and use additional languages. We also want to introduce user profiles in order to provide user-adapted recommendations. Finally, we want to adapt the developed App for iOS-based mobile devices.

Acknowledgements This work was supported in part by Projects MINECO TEC2012-37832-C02-01, CICYT TEC2011-28626-C02-02, CAM CONTEXTS (S2009/TIC-1485).

References

1. Akcayir, M., Akcayir, G., Pektas, H., Ocak, M.: Augmented reality in science laboratories: The effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. Computers in Human Behavior **57**, 334–342 (2016)
2. Griol, D., Callejas, Z, López-Cózar, R., Riccardi, G.: A domain-independent statistical methodology for dialog management in spoken dialog systems. Computer, Speech and Language **28**(3), 743–768 (2014)
3. Hansmann, U., Merk, L., Nicklous, M., Stober, T.: Pervasive computing: The mobile world. Springer (2003)
4. Klopfer, E., Squire, K.: Environmental detectives - the development of an augmented reality platform for environmental simulations. Educational Technology Research and Development **56**(2), 203–228 (2008)
5. López-Cózar, R., Callejas, Z.: Multimodal dialogue for ambient intelligence and smart environments. In: Handbook of Ambient Intelligence and Smart Environments, pp. 559–579. Springer (2010)
6. McTear, M., Callejas, Z.: Voice Application Development for Android. Packt Publishing (2013)
7. Olsson, T.: User Expectations and Experiences of Mobile Augmented Reality Services. Ph.D. thesis, Tampere University of Technology (2012)
8. Pieraccini, R.: The Voice in the Machine: Building computers that understand speech. MIT Press (2012)
9. Polvi, J., Taketomi, T., Yamamoto, G., Dey, A., Sandor, C., Kato, H.: SlidAR: A 3D positioning method for SLAM-based handheld augmented reality. Computers and Graphics **55**, 33–43 (2016)
10. Tapia, D., Abraham, A., Corchado, J., Alonso, R.: Agents and ambient intelligence: case studies. Journal of Ambient Intelligence and Humanized Computing **1**(2), 85–93 (2009)

Creating Virtual Humans with Game Engines for Evaluate Ambient Assisted Living Scenarios

Manuel Sánchez Palacios, Juan Antonio Álvarez-García, Luis Miguel Soria and Damián Fernández Cerero

Abstract In order to evaluate AAL systems, virtual environments help to reduce costs and time, but these environments do not include hyper-realistic human movements. This is something crucial to evaluate activity recognition systems. The present work in progress describes how it is looking for a way to solve this problem and its development using a virtual environment. By means of a game engine and adjusting their parameters, simulations of real acceleration data sets have been generated. It is continuing looking for a valid model to follow.

1 Introduction

Ambient Assisted Living (AAL) is a research area that uses technology to improve the quality of life of elderlies, by increasing their autonomy in daily activities, and by enabling them to feel secure, protected and supported. In order to ensure the effectiveness and usability of AAL approaches, leading to an eventual real-life application, researchers must agree on a set of standard evaluation methods. However, this is a difficult task because of the diversity of both the functions performed by AAL solutions, as well as the heterogeneity of sensors and other hardware used in this field. A function that has emerged as central to many AAL solutions is Activity Recognition (AR), mainly because the ability to understand the user is situation and context are key for real-life usability. Evaluation of AR Systems (ARS) initially was carried out on specific datasets to each system, recorded in laboratory settings, and they generally achieved high accuracy. However, since each research group evaluated their system using their own dataset, a comparison between different ARS was almost impossible. The problem is even worse when the subjects under study are elderly. Most of the current works focused on this section of

M.S. Palacios · J.A. Álvarez-García(✉) · L.M. Soria · D.F. Cerero
Computer Languages and Systems Department, University of Seville, 41012 Seville, Spain
e-mail: {msanpal,jaalvarez,lsoria,damiancerero}@us.es

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_12

the population, make use of synthetic dataset obtaining from younger users imitating their behaviour. The first improvement towards standardized test was achieved when the first benchmark datasets were developed, enabling researchers to compare different ARS. However, for ARS, this approach is not enough because the comparison is limited to systems that use the same configuration, type and placement of sensors, as the one used while recording the dataset.

To tackle this challenging issue, live competitions appeared. The first living competition in ARS was EvAAL-AR [1, 2] (Evaluating AAL Systems through Competitive Benchmarking − Activity Recognition). In these contests, each team is required to bring their own ARS, which is evaluated using criteria in the following features: recognition accuracy, user acceptance, recognition delay, installation complexity, and interoperability with AAL systems. Although the EvAAL-AR competition has been an interesting experience, several issues appeared, including reduced number of participants and longer test desired.

These points focus our attention on virtual environments, where a most extensive evaluation can be carried out. In this paper, we describe the existing virtual environments focused on AAL with humans.

The rest of the paper is organized as follows: Section 2 gives an overview of the state of the art. The work in progress of a game engine with hyper realistic human movements is presented in section 3. Section 4 draws the actual conclusions and next steps.

2 Related Work

The testing and the validation of AAL software systems in real environments require huge efforts in terms of work, time and money. Virtual AAL environments allow simulation of humans with wearable devices, virtual sensors in the environment and, ideally, a physic exact to the real world. They will be an effective solution to speed up the development and the testing phase of AAL systems, by limiting the problems that could arise during the tests in real scenarios. They also tackle the problems of large periods of simulation and multiple people in the environment reducing costs and time.

Some tools are designed for simulating real world, but normally are specific for robotics. In this latter, the movements are generated from automatic or human controlled actions. Humans can be included in these simulators, but their movements are not realistic at all[1]. In the next section, AAL simulators where humans can be included are described.

2.1 AAL Simulators

AAL simulators are not a new concept, Ubiwise[4] in 2002 was one of first approaches. It focuses on the use and analysis of environment models for ubiquitous

[1] https://vimeo.com/90004660

computing systems. People are considered individually on a simulation engine based on Quake II, where real people, acting as players in a game, generate information about their own context, which is captured through simulated sensors.

Another interesting tool is UbiREAL[5]. It lets users intuitively grasp how devices are controlled, depending on temporal variation of contexts in a virtual space. The main contribution of UbiREAL is that it simulates physical quantities (e.g. temperature or humidity) and includes a network simulator. Their human movements are not realistic[2].

Roalter, et al. [6] presented a tool chain to simplify research on Intelligent Environments. The tool chain consists of: a middleware (Robot Operating System tool or ROS), environment edition (SweetHome 3D tool), and simulation and visualization (Gazebo tool). The ROS middleware plays an important role in making interactions possible.

In [7] an AAL virtual simulator is developed build on the top of Modular Robots Open Simulation Engine (MORSE[3]). The authors recreate a virtual home automation environment where a handicapped user moves using a wheelchair. The environment includes a virtual thermometer, RFID sensors on the doors, gas detectors, flowmeters and heart rate and apnea detectors. The application on top of this home automation environment is a virtual caregiver that can detect emergency scenarios. This approach is close to our work but we focus on human movement realism and in the scenario proposed by the authors we cannot check the realism of the human movement (only the one of the wheelchair)[4].

Campillo-Sanchez, et al. work [8,9] advances with Ubik Mobile and PHAT architectures to provide developers with tools to test context-aware services based on smartphones in a simulated AAL environment. Their main contribution is that a real smartphone, and not only an emulator, can be connected to the simulated world to conduct manual and automatic tests. Furthermore PHAT combines a game engine, jMonkeyEngine, with a Multi-Agent Based simulator(MaSON) obtaining realistic human movements and behavior.

From biomechanical research [10] there are specific tools such as AnyBody™[5] that model virtual humans very detailed[6] and include motion capture techniques but normally it does not focus on AAL environments so it is out of this scope.

2.2 Virtual Sensors

Virtual sensors, as opposed to physical sensors, can provide indirect measurements by using other available sensor data, models and knowledge. Raveendranathan et al. [11] proposed a virtual sensor framework in the gait analysis domain for the purpose of enabling real-time activity and posture

2 https://www.youtube.com/watch?v=57juhBRbIsY
3 http://www.openrobots.org/morse/doc/stable/morse.html
4 https://www.youtube.com/watch?v=IBuYZFoKmy8
5 http://www.anybodytech.com/index.php?id=publications
6 https://www.youtube.com/watch?v=SmJRkjgpoSg

recognition. vanBeek et al. [12] simulated the time trial of athletes in the Tour de France with various physiological and mechanical models.

Our approach initially is simpler, using physics engine from Unity to create a virtual accelerometer although to more complex sensors another option is to include an Android emulator that includes an accelerometer, gyroscope and magnetometer [9].

3 Tools and Development

Our main goal is to develop an AAL simulator that allows hyperrealist human movements based on real-world physics using a game engine. Once this objective is achieved, measuring acceleration in any part of the model is body (wrist, chest, hip, etc.) will be possible using virtual-wearable sensors. This work in progress describes the steps to achieve this aim. In brief, next stages are involved to develop the virtual environment ecosystem:

1. Select a game engine.
2. Modify the physics engine to obtain real-world physics.
3. Generate a virtual human with enough animation to get hyperrealism.
4. Compare the measurements of our hybrid physics+animation engine with real datasets.

Steps 2 to 4 have been repeated to achieve good results.

The first step was critical to get our objective. Although open source game engines such as JMonkeyEngine, allow the modification of all the code, most popular game engines (Unity, Unreal Engine 4, Source 2 and CryENGINE) are more accepted, their developer community is much greater and introducing latest technical advances quicker. Finally, Unity, a well-known gameengine oriented to 3D games, was selected to model the physics in the AAL environment. These specifications make it a powerful tool to develop the first testing scenario. One of the main advantages of this engine is the possibility to import and export existing formats in a large number of platforms. This will help to share the test environment with other platforms. Furthermore, Unity is the leader in developing gaming platforms, with more than four millions registered developers and it is used in more than 45% of entire market projects for game engines [7], about three times the rate of the nearest competitor.

Developers might encounter issues with physics in Unity, mainly because this game engine is not oriented to real world simulations. By hence, it does not use a metric system. To attach the engine with real world physics units, some adjustments have been performed. It is assumed that 1 unit in Unity is 1 meter in the real environment and the gravity is modified to approach to 9.81 m/s^2. Friction is needed to decrement movement, but Unity is not prepared for ground reaction force, so a modification on the behaviour of the physics materials has been included.

The third point refers to the process of achieving the hyperrealism in humans.In this vein, a mixed model of physics+animation was chosen. The combination of physics and animations is necessary because Unity's physic engine does not include fine-grained movements. By other side, using animation is not suitable to simulate real physics and get distinct reaction in different instances with the same path. The more promising model for animations is Mixamo Test Model [8]. The Mixamo default animations are not as realistic as they should, but with some modifications and simplifications as we can see in Figure 1 and adding a Rag Doll blended with animation the result should be better.

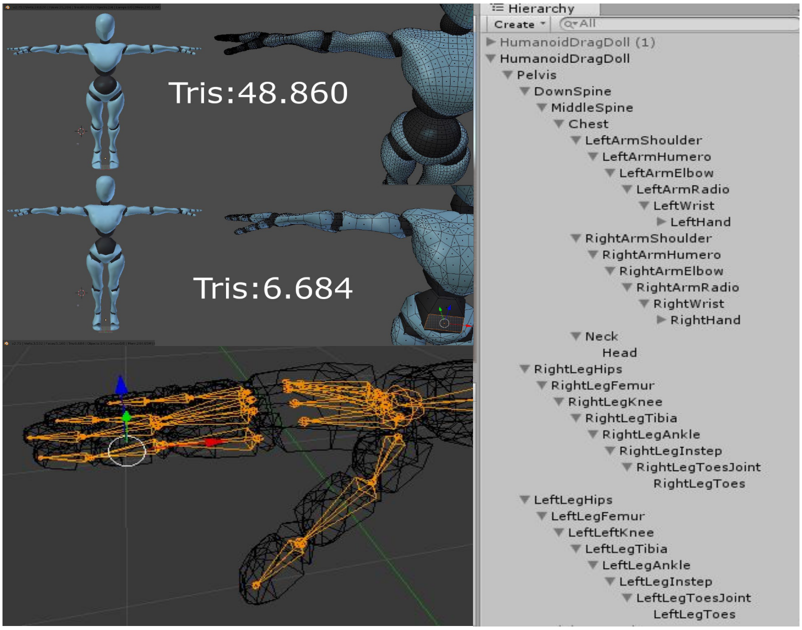


Fig. 1 Simplifying Mixamo model (top left), Including new bones (bottom left), Hierarchy of human Rag Doll (right)

The last step checks how close are the virtualized human actions to real-world physics movements. In order to evaluate the progress, a virtual accelerometer to our character was placed on the scene, performing movements included in a real dataset. The dataset used in this test was UCI Donald Bren [9], which collects accelerometer data from an Android Phone carried on the chest as a sensor. To compare distinct experiments with the dataset, there was established some Key Performance Indicators (KPIs), indicating the accuracy of the reading from the virtual environment compared to the real dataset. Iterating with these KPIs, the current virtual environment approaches to the real data. The KPIs values used were: Maximum acceleration, Minimum acceleration, Curve behaviour, Average acceleration and Histogram of Frequency.

These KPIS are the first minimum KPIs if the model does not get good results with these key numbers we can discard or iterate the model. In the future new more accurate KPIs will be defined to check the data but in the actual state of the project we are checking the environments created viability and the checking is continued during today.

Iterating only an animations environment, the results of the KPIs were better. Once the results were collected, the KPIs were too fixed for a real world environment but some parameters were well-adjusted. The KPIs were acquired again in an absolute-physical model. However, iterating in this profile, the results were unsatisfactory. Finally, now the system is testing with a mixed environment profile, where animations were more accurate aided by real physics. When the final configuration was reached and good KPIs were achieved, a dataset builder with Android smartphones as sensor was developed to ensure that the paths followed by the virtual agents, and the virtualized test are as accurate as the real experiment is.

As we see in the Figure 2 with the same frequency we have much more iteration and noises that in the real sample does not appear. These graphics are extracted by the only animation environment. Here an emulated accelerometer chest sensor where comparing it is KPIs with the UCI data sets. In this environment adjusting the animations good KPIs could be finally reached but every time the same test is reproduced, exactly the same data will be obtained and it is not a real world approach, but is a good beginning for closed and repeated test, which this investigation were not looking for.

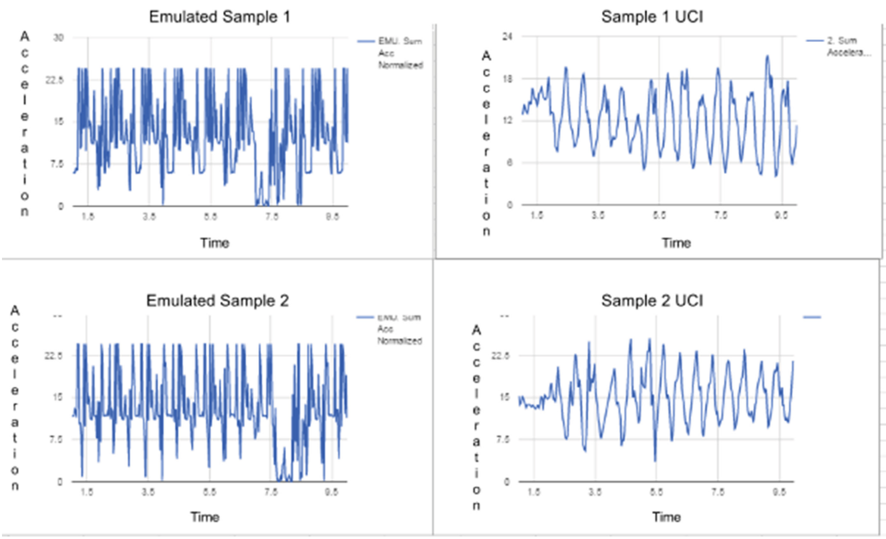


Fig. 2 Comparing UCI Dataset with simulation results.

There are several software tools to model, create, render and execute experiments in real time. We considered SketchUp to model the Living Lab, it is the most arquitectonic one, although there are other options such as SweetHome 3D or Autodesk Homestyler. It allows to easily create models of domestic environments, in which the various sensors and actuators can be placed.

For modelling, rigging and animating the agents, it was used Blender which is a very powerful free software easy to use and with professional performance.

4 Conclusions

In this work in progress, we explained the current trend to use virtual environments to test real solutions in AAL, analysed the state of the art of AAL simulators including humans with realistic movements and showed the steps given to achieve our goal: measure acceleration with wearable virtual devices in any part of our virtual humans.

As future work, we want to include realistic human behaviour such as [13] and create a framework that allows interaction with external researchers such as [14]. This kind of framework could ease focus on specific goals such as [15].

Acknowledgments This research is partially supported by the projects of the Spanish Ministry of Economy and Competitiveness, HERMES (TIN2013-46801-C4-1-r) and Simon (TIC-8052) of the Andalusian Regional Ministry of Economy, Innovation and Science.

References

1. Álvarez-García, J.A., et al.: Evaluation of localization and activity recognition systems for ambient assisted living: The experience of the 2012 EvAAL competition. Journal of Ambient Intelligence and Smart Environments **5**(1), 119–132 (2013)
2. Gjoreski, H., et al.: Competitive Live Evaluations of Activity-Recognition Systems. Pervasive Computing, IEEE **14**(1), 70–77 (2015)
3. Álvarez-García, J.A., et al.: Evaluating wearable activity recognition and fall detection systems. In: 6th European Conference of the International Federation for Medical and Biological Engineering. Springer International Publishing (2015)
4. Barton, J.J., Vijayaraghavan, V.: Ubiwise, a ubiquitous wireless infrastructure simulation environment. HP Labs (2002)
5. Nishikawa, H., et al.: UbiREAL: realistic smartspace simulator for systematic testing. In: UbiComp 2006: Ubiquitous Computing, pp. 459–476. Springer, Heidelberg (2006)
6. Roalter, L., et al.: Developing intelligent environments: a development tool chain for creation, testing and simulation of smart and intelligent environments. In: 2011 7th International Conference on Intelligent Environments (IE). IEEE (2011)
7. Calvaresi, P., et al.: 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)

8. Campillo-Sanchez, P., Gómez-Sanz, J.J., Botía, J.A.: PHAT: physical human activity tester. In: Hybrid Artificial Intelligent Systems, pp. 41–50. Springer, Heidelberg (2013)
9. Campillo-Sanchez, P., Serrano, E., Botía, J.A.: Testing context-aware services based on smartphones by agent based social simulation. Journal of Ambient Intelligence and Smart Environments **5**(3), 311–330 (2013)
10. van den Bogert, A.J., et al.: A real-time system for biomechanical analysis of human movement and muscle function. Medical & Biological Engineering & Computing **51**(10), 1069–1077 (2013)
11. Raveendranathan, N., et al.: From modeling to implementation of virtual sensors in body sensor networks. Sensors Journal, IEEE **12**(3), 583–593 (2012)
12. Van Beek, J.H.G.M., et al.: Simulating the physiology of athletes during endurance sports events: modelling human energy conversion and metabolism. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **369**(1954), 4295–4315 (2011)
13. Li, Y., Guo, Y.: Wiki-Health: from quantified self to self-understanding. Future Generation Computer Systems **56**, 333–359 (2016)
14. Campuzano, F., et al.: Generation of human computational models with machine learning. Information Sciences **293**, 97–114 (2015)
15. Serrano, E., Poveda, G., Garijo, M.: Towards a holistic framework for the evaluation of emergency plans in indoor environments. Sensors **14**(3), 4513–4535 (2014)

Metabolic.Care: A Novel Solution Based on a Thermography for Detection of Diabetic Foot

Virginie Felizardo, Hugo Rodrigues, Nuno C. Garcia, Celina Alexandre, Daniel Oliveira, Paula Sousa, Nuno M. Garcia and Nuno Pombo

Abstract The use of computer supported diagnosis for the early detection of illness-related complications is of utmost importance not only for the ill individual, but also for the society in general. This paper describes one approach to address the early detection of complications in diabetic feet, achieved by the use of a new device created using out-of-the box components, and complemented with a software ecosystem. This paper describes the solution and the ongoing research.

Keywords Diabetes Mellitus · Diabetes mellitus monitoring · Thermography · Clinical image processing

---

V. Felizardo · H. Rodrigues · N.C. Garcia · C. Alexandre · D. Oliveira · P. Sousa · N.M. Garcia · N. Pombo(✉)
ALLab Ambient Living Computing and Telecommunications Laboratory, Instituto de Telecomunicações, Universidade da Beira Interior, Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal
e-mail: {virginie,hrodrigues,ncgarcia,celinaalexandre,danieloliveira,paulasousa}@it.ubi.pt, ngarcia@di.ubi.pt, ngpombo@ubi.pt

V. Felizardo · C. Alexandre · D. Oliveira · N.M. Garcia · N. Pombo
Computer Science Department, Faculty of Engineering, University of Beira Interior, Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal

N.M. Garcia
School for Communication, Architecture, Art and Information Technologies (ECATI), Lusophone University of Humanities and Technologies (ULHT), Campo Grande 376, 1749-024 Lisbon, Portugal

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_13

1 Introduction

Unmonitored and untreated Diabetes mellitus (DM) may cause not only lower limb amputation as a result of the ulceration of the feet, but also many microvascular complications, leading to high morbidity and mortality. In addition, as many as 50% of people with DM are prone to develop long-term complications as a peripheral neuropathy (PN) (1). The foot-specific effects of PN are principle factors in the development of deformity, elevated plantar pressures and the increased risk for plantar ulceration (2). Moreover, the loss of intrinsic foot muscle function in people with PN is of particular concern because muscle atrophy is one of the earliest detectable precursors of abnormal foot function leading to pathology (3), whose consequences include foot muscles attenuation, and a reverse, longitudinal arch deformation under increasingly loaded conditions (4).

Therefore, it is important to define and control modifiable risk factors that contribute to diabetic complications. In line with this, there are promising computerized systems which enable the assessment, the prediction and/or the monitoring of patients suffering with DM. Additionally, it is estimated, according to information from the International Diabetes Federation that 85% of lower limb amputations (5), could be reduced with eHealth systems and techniques for the early detection of diabetic foot. Thus, the Metabolic.Care as a subproject (products, processes or systems workgroup 10 - PPS # 10) of TICE. Healthy project (2011-2014) (6), aims to develop a monitoring and feedback solution for patients with diabetes who presented symptoms associated with diabetic foot, with full awareness that these scenarios represent a growing concern in society as a result of changes in lifestyles and alimentation.

This paper describes the process flow for the use of a new device created in the context of a funded project, starting with the identification of the problem, and ending with foreseen future work. The remainder of this chapter is organized as follows: this paragraph concludes section 1, the Introduction; section 2 addresses in detail the implemented solution; section 3 presents results of the comprehensive experimental evaluation; and finally, section 4 discusses the meaning and significance of the results and concludes the chapter.

2 Methods

The proposed solution comprises a physical equipment (thermographic image acquisition equipment) and an application hosted in the web platform, eVida (5). This platform, as presented in Figure 1, named Metabolic.Care Thermal Scanner (MCTS), is responsible for collecting thermographic images on a conventional query scenario, obtains images through digitalization of liquid crystal sheets after contact with the patient's feet.

The MCTS is made of out-of-the-box components, mounted on a stainless steel frame for durability and weight resistance. The main component of the MCTS is an A3 standard page scanner. On top of the scanner, a set of thermo-sensitive liquid crystal sheets was adapted. The sheets are also a easily found commodity, and the main criteria for its selection was its temperature range, as it is intended that these sheets also prove to be sensitive to the temperature range of the feet and its special thermic resolution. To select the appropriate sheets, several thermo-sensitive liquid crystal sheets were calibrated and its thermic resolution in space was assessed, as to guarantee a minimum level of contamination in space, and therefore, a great level of detail in terms of variation of temperature per square millimeter.

Fig. 1 Thermo imaging scanner for feet.

As Metabolic.Care aims at providing complementary technological means for early diagnosis of complications of the diabetic foot condition, by enabling a greater connection between patients and clinicians and allowing closer monitoring of patients through the use of ICT, a set of applications was also developed to address the capture, storage, transmission and analysis of the images. The ultimate purpose of the project is the adoption of this platform to use in a clinical setting for data analysis and detection of early symptoms.

In line with this, this solution includes the capability to transmit information to the central platform eVida, using a standard for medical images such as Digital Imaging and Communications in Medicine (DICOM) (7). After the capture of the images, and before or after its transmission to the central repository of data, a software is used to detect asymmetry in the images of the feet, triggering alarms either when an asymmetry is detected and when an negative evolution of the images is found. In fact, the system allows sending pictures and accessing online clinical records and/or progression of the disease, ideally, combined with a friendly user interface.

3 Results

As depicted in Figure 2, the workflow of the proposed system is presented as follows. Firstly, on a clinical visit, the patient is asked to be seated while barefoot. Then, he/she places his/her feet on the above mentioned device in order to acquire his thermal feet image. It was estimated that the feet must not sensitize the liquid crystal sheets for more than 2 minutes, otherwise, there is too much contamination from neighbouring heat sources in the feet, and the resulting image loses detail. It is also worth mentioning that this procedure occurs after the user has his/her feet exposed at the room's temperature for about 2 to 5 minutes, to allow the feet temperature to stabilize. Secondly, the physician briefly checks the image and proceeds to upload the image to the hosted application on the eVida platform, inserting relevant information acquired during the appointment. The physician can select to collect multiple images, as there is no need to change the liquid crystal sheets between image collection procedures. To allow for proper sanitation, the use of a very thin transparent plastic film over the liquid crystal sheets is advised.

Thirdly, the uploaded image is processed, beginning with the segmentation of the feet, allowing the resulting segments to be compared one with another using the Grabcut method (8) as presented in Figure 3. This method is based in cuts on the image, beginning with a predefined bounding box around the object to be segmented, calculating the distribution of colors between the object and the background through a Gauss model which results on a construction of a random field of pixels (9). That along with an energy function optimizes the graphical cut.

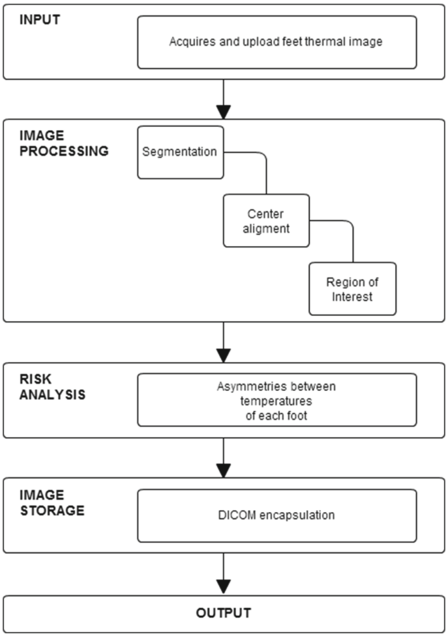


Fig. 2 System workflow.

Fourthly, the segmented images are rotated in order to be aligned with the center and then both segments are split in Regions of Interest (ROI) which are compared so as to determine the risk of developing problematic conditions for diabetic foot. This risk is obtained from asymmetries between temperatures of each foot (comparing the color distance from each pixel). When the distance is greater than a defined threshold, it means that there might be some problem, and therefore an alert is raised for the physician at the software and the database. Finally, both uploaded image and the evaluation result are encapsulated in the DICOM format, and stored in the eVida platform. All the images are accessed using an authenticated user Metabolic.Care profile on the eVida platform. Depending the users' profile, some activities are allowed such as: load and process images, consult clinical records, delete images and write and/or edit comments. Both patients and health care professionals can interact with the platform. However, patients have a limited access which is restricted to the consultation of clinical history, and to load new images.

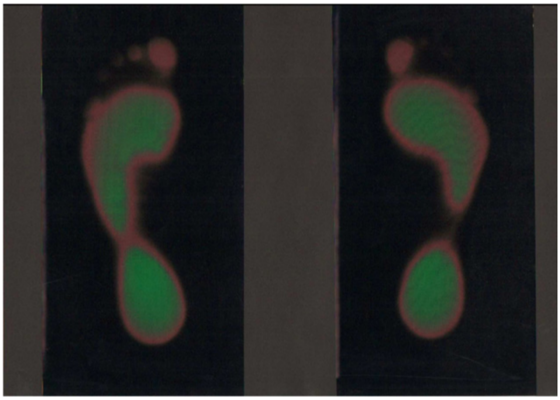


Fig. 3 Sample image of left and right healthy feet as captured by the platform.

4 Discussion and Conclusions

Early diagnose of diseases is rightfully viewed as both a strategy and the best therapeutic. In the case of *Diabetes Mellitus*, the lesions that this disease can bring to the patients' feet are not only severe because it jeopardize their ability to move, work and live a regular life, to extremes, it can be life threatening. The software that interacts with the platform allows the clinician, to not only diagnose on the spot possible lesions, but also to record thermography images, whose impact in the development of future diagnosis is still to be evaluated, in a much more automated procedure when compared with the standard tests. Although there are many research works on the use of thermography as a diagnostic method to early identify potential signs of the diabetic feet, currently there are no cheap tools to support this diagnostic procedure. In addition, the DICOM format offers to the platform a desirable interoperable data which may leads to a low-cost, inter-connectable and shareable system. Moreover, the proposed system allows the HL7 standard. Both DICOM and HL7 data are stored using a Resource Information Model (RIM) which provides a local URL to the compliant repository. The communication with external applications is supported by the Mirth Connect broker (10). Despite the platform and its supporting software having been build, calibrated and tested, there is still a currently ongoing work of further validation of the platform. These activities include the use of the “golden standard” methods in the early assessment of the diabetic foot condition, e.g. the 10 g monofilament test that is a sensory test to evaluate loss of protective sensation of foot, the 128 Hz diapason to evaluate the vibration sensation, an oximetry measurement applied to feet fingers, thermographic still images captured with a camera, and finally, standard digital photography and video. Additionally, the thermographic images of the user's’ feet are also captured using the Metabolic.Care. There is also the goal of making this data publicly available at the ALLab laboratory mediawiki (to be found in https://allab.it.ubi.pt/mediawiki). Finally, as future work, and also a manner to circumvent the cases where there no possibility to use the asymmetry on the color map for the two feet, either because one of the feet is already diagnosed and shows lesions, or because the user has only one foot, the application will also keep track of image measurements through time, being in this case the comparisons made not only between two feet, but also for each foot, from its current state in comparison with previously recorded states.

Acknowledgements This project was co-funded by the QREN, COMPETE and European Union as project no. 13842 – TICE.Healthy - Systems of Health and Quality of Life. The authors wish to thank the opportunity and financial support that permit to carry on this project. The authors would also like to acknowledge the contribution of the COST Action IC1303 – Architectures, Algorithms and Platforms for Enhanced Living Environments (AAPELE). Contributing to this research, the authors affiliated with the *Instituto de Telecomunicações* also acknowledge the funding for the research by means of the program FCT project UID/EEA/50008/2013. (Este trabalho foi suportado pelo projecto FCT UID/EEA/50008/2013). Finally, the authors would like to show their appreciation to all the TICE team, in particular to the TICE.Metabolic.Care team at the University of Coimbra and at Exatronic, Aveiro, Portugal, for the fruitful discussions and the permission to use some images.

References

1. Gordois, A., Scuffham, P., Shearer, A., Oglesby, A., Tobian, J.A.: The Health Care Costs of Diabetic Peripheral Neuropathy in the US. Diabetes Care **26**(6), 1790–1795 (2003)
2. Cheuy, V.A., Hastings, M.K., Commean, P.K., Ward, S.R., Mueller, M.J.: Intrinsic foot muscle deterioration is associated with metatarsophalangeal joint angle in people with diabetes and neuropathy. Clinical Biomechanics **28**(9–10), 1055–1060 (2013)
3. Greenman, R.L., Khaodhiar, L., Lima, C., Dinh, T., Giurini, J.M., Veves, A.: Foot Small Muscle Atrophy Is Present Before the Detection of Clinical Neuropathy. Diabetes Care **28**(6), 1425–1430 (2005)

4. Kelly, L.A, Cresswell, A.G, Racinais, S., Whiteley, R., Lichtwark, G.: Intrinsic foot muscles have the capacity to control deformation of the longitudinal arch. Journal of the Royal Society Interface [Internet] **11**(93), January 29, 2014. http://rsif.royalsocietypublishing.org/content/11/93/20131188.abstract
5. IDF [Internet]. http://www.idf.org/
6. TICE [Internet]. https://www.tice.pt/
7. Mustra, M., Delac, K., Grgic, M.: Overview of the DICOM standard. In: ELMAR, 2008 50th International Symposium, pp. 39–44 (2008)
8. Gao, Z., Shi, P., Karimi, H.R., Pei, Z.: A mutual GrabCut method to solve co-segmentation. EURASIP Journal on Image and Video Processing **2013**(1), 1–11 (2013)
9. Kohli, P., Osokin, A., Jegelka, S.: A principled deep random field model for image segmentation. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1971–1978 (2013)
10. Mirth Connect [Internet]. http://www.mirthcorp.com/products/mirth-connect

Identification of Activities of Daily Living Using Sensors Available in off-the-shelf Mobile Devices: Research and Hypothesis

Ivan Miguel Pires, Nuno M. Garcia, Nuno Pombo and Francisco Flórez-Revuelta

Abstract This paper presents a PhD project related to the identification of a set of Activities of Daily Living (ADLs) using different techniques applied to the sensors available in off-the-shelf mobile devices. This project consists on the creation of new methodologies, to identify ADLs, and to present some concepts, such as definition of the set of ADLs relevant to be identified, the mobile device as a multi-sensor system, review of the best techniques for data acquisition, data processing, data validation, data imputation, and data fusion processes, and creation of the methods for the identification of ADLs with data mining, pattern recognition and/or machine learning techniques. However, mobile devices present several limitations, therefore techniques at each stage have to be adapted. As result

I.M. Pires(✉) · N.M. Garcia · N. Pombo
Instituto de Telecomunicações, University of Beira Interior, Covilhã, Portugal
e-mail: impires@it.ubi.pt, {ngarcia,ngpombo}@di.ubi.pt

I.M. Pires
Altranportugal, Lisbon, Portugal

I.M. Pires · N.M. Garcia · N. Pombo
ALLab - Assisted Living Computing and Telecommunications Laboratory,
Department of Informatics, University of Beira Interior, Covilhã, Portugal

N.M. Garcia
ECATI, Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal

F. Flórez-Revuelta
Faculty of Science, Engineering and Computing, Kingston University,
Kingston upon Thames, UK
e-mail: F.Florez@kingston.ac.uk

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_14

of this study, we presented a brief review of the state-of-the-art related to the several parts of a mobile-system for the identification of the ADLs. Currently, the main focus consists on the study for the creation of a new method based on the analysis of audio fingerprinting samples in some Ambient Assisted Living (AAL) scenarios.

Keywords Sensors · Data fusion · Mobile devices · Activities of Daily Living · Data acquisition · Data processing · Data imputation · Audio fingerprinting · Pattern recognition

1 Introduction

Mobile devices, *e.g.,* smartphones, tablets, smartwatches, and other specific devices, are commonly used while performing Activities of Daily Living (ADLs). Mobile devices may be considered in Ambient Assisted Living (AAL) systems, because they are able to collect data related to the user's environment and user activities, which can be used for different purposes, including the identification of ADLs, what is important to support the independent living or the management of chronic condition of elderly people and/or individuals with some type of disability.

This study is part of a PhD project that aims to create a new method to recognize ADLs with good accuracy, and to reduce the constraints of mobile devices for these tasks. This project started with the definition of two different sets: ADLs that may be reliably identified with mobile devices, and, valid sensors available in mobile devices. This work continues the research presented in [1], which included the calculation of a jump flight time that makes the use of pattern recognition techniques to identify patterns in vertical jumps [2]. Another work where similar analysis has been carried out is related to the Heel-Rise test [3], based on a test used mainly by physiotherapists that allows the detection of fatigue and/or specific diseases. A preliminary study for the identification of ADLs using mobile devices was already presented in [4], which will pave the way for the development of a personal digital life coach [5].

The proposed method includes three different stages, as follows: acquisition and processing, data fusion, and identification of ADLs based on the fused data. The first stage includes the activities of data acquisition, data processing, data validation, data cleaning, and data imputation. The second stage is focused on data fusion, which consists on the consolidation of the data acquired from all sensors, whereas the last stage consists on several techniques, including data mining, pattern recognition and machine learning techniques combined with the users' feedback in order to identify the ADLs. The main goal is that the method could provide an user profile based on events inferred from a regular activity day, making use of multi-sensor data fusion technologies and context awareness approaches, combined with other information available from the user context.

The remaining sections of this paper are organized as follows. Section 2 presents a review of the state of the art, focusing in the main concepts of this work. Section 3 introduces the proposed method to be developed during the PhD project, presenting a small overview of the mobile multi-sensor system, methods that can be applied to sensors available on off-the-shelf devices, the current approaches obtained and the work in progress. Section 4 presents the discussion and conclusions.

2 Related Work

This research involves several topics, such as the identification of ADLs, the sensors available in off-the-shelf mobile devices and the different methods for acquisition, processing, validation, imputation, cleaning, and fusion of the data acquired. To conclude, a summary of the review about data mining, pattern recognition and machine learning techniques are presented.

2.1 Identification of ADLs

People's self-care activities are commonly named as ADLs, which involve a desire to achieve a degree of physical comfort, self-care, and autonomy, that may promote feelings of independence and personal control. Examples of ADLs include activities related to personal appearance and hygiene, domestic skills, household management, family and child care, conversational and social skills, and leisure, education, training and work activities [6].

Several sensors, including the accelerometer, gyroscope, magnetometer, communication, proximity, light, gravity, RFID (Radio-Frequency IDentification) sensors, GPS (Global Positioning System) receiver, microphone and camera, may be used to detect when a person is having/preparing a meal, washing up, bathing, waking up, sleeping, standing, sitting, watching TV, using the phone, doing chores, cycling, jogging or perform other activities [7-10].

2.2 Sensors

Sensors are small hardware components that capture different types of signals, which are widely available in several mobile devices, including smartphones, smartwatches, and tablets, which may be used to collect data in a plethora of situations, including the identification of ADLs [11].

Sensors may be used for different purposes and working environments, collecting different types of data [12]. On the one hand, these environments can be classified as controlled, uncontrolled, static, dynamic, uncertain and undefined. On the other hand, sensors, can be categorized in mechanical, environmental, motion, imaging, proximity, acoustic, medical, chemical, optical, and force sensors.

Mobile devices may be considered as a multi-sensor platform, because they are equipped with several types of sensors whose combination may increase the reliability in the detection and identification of ADLs. However, the most important stage in multi-sensor systems is the signal classification with pattern recognition or machine learning methods [13]. One example is a wearable multi-sensor ensemble classifier for physical activity pattern recognition, developed in [14], which combines several classifiers based on different sensor feature sets to improve the accuracy of physical activities identification and recognition.

2.3 Data Acquisition

Data acquisition is supported with the sensors available in off-the-shelf mobile devices, which allows anywhere and at anytime the capture of data and conversion of the electrical signals received by each sensor into a readable format [15]. Several challenges are associated with the data acquisition process when recognizing ADLs, including the positioning of the mobile device, the data sampling rate and the number of sensors to be managed [16]. This process can be enhanced with several frameworks, including the *Acquisition Cost-Aware QUery Adaptation (ACQUA)*, consisting in a query processing engine implemented for mobile devices that dynamically modifies both the order and the segments of data streams requested from different sources [17].

2.4 Data Processing

Data processing may be implemented by two different architectures, such as the Device Data Processing Architecture and the Server Data Processing Architecture [18]. The first one, is designed to acquire the data from the sensors embedded in a mobile device and process the data locally. This architecture is useful when the processing methods require low resources, such as processing the accelerometer data, proximity sensor data, and others. On the contrary, Server Data Processing Architecture consists in the dispatch of the data collected to a remote server allowing the computation of a larger amount of data as well as computations of complex nature. This architecture is useful to work with data acquired by a GPS receiver or imaging sensors.

2.5 Data Validation

Data validation is the process executed simultaneously with data acquisition and data processing in order to validate if the data are correctly acquired or processed so as to obtain valid, accurate and reliable results. During the Data Validation stage, some tests are performed depending on the characteristics of the application, *e.g.* if it depends on continuous data collected during a time interval, if any value at a random instant is not correctly acquired or missing, the value should be inferred using data imputation techniques. Aiming at reducing noise artifacts the data captured or processed should be cleaned, treated or imputed, and therefore more prone to offer better and reliable results.

Data validation methods are mainly implemented with machine learning algorithms, such as principal component analysis (PCA), relevance vector machine (RVM) and artificial neural networks (ANN) [19, 20].

2.6 Data Imputation

Data imputation is applied when the acquired data is faulty. The missing data failures can be classified as Missing Completely At Random (MCAR), Missing At Random (MAR) and Missing Not At Random (MNAR) [21]. MCAR happens when missing values are randomly distributed across all observations. MAR is observed when missing values are randomly distributed within one or more subsamples instead of the whole data set. MNAR is the type of missingness that arises when missing values are not randomly distributed across observations. Several methods for data imputation are presented in [22], being the K-nearest neighbors (k-NN) the most used method [23].

2.7 Data Cleaning

Data cleaning is process similar to data imputation and is executed when the data contains incorrect values, also considered as noise. Data cleaning methods may be statistical and probabilistic methods, which are used to apply filters to the collected data and adjust the values into correct measurements [24]. The data cleaning process may be supported on spatial and temporal characteristics of the collected data.

2.8 Data Fusion

Data fusion is the most important step for the integration of the data collected by several sources, aiming at to increase the reliability of the algorithms for the identification of ADLs with mobile devices, to reduce the effects of the incorrect data captured by the sensors, or to support the data processing [25].

Data fusion methods may be executed either on the mobile application, as a background process, or on a remote system. These methods may be categorized as probabilistic, statistic, knowledge base theory and evidence reasoning methods [26]. Although, the most used method for data fusion is the Kalman filter, which is a dynamically weighted recursive least-squares algorithm [27].

2.9 Data Mining, Pattern Recognition and Machine Learning Techniques

This is the last step in the recognition of ADLs, comprising the identification of the main features of the collected data, which may include the peaks of acceleration, the standard deviation of the acceleration, the maximum and minimum values of acceleration, the amplitude of the audio data captured, and others explored during this PhD project.

The use of data mining, pattern recognition and machine learning techniques allows the classification of the sensors' data. Several studies have been carried out to identify ADLs [28-35].

3 Proposed Solution

The proposed solution in this PhD project consists on a multi-sensor mobile system in which several techniques are applied aiming to create new methods to identify ADLs with high accuracy.

3.1 Method Description

The proposed method consists on an architecture to support the development of applications for the monitoring and identification of ADLs. Mobile devices combined with mobile applications are able to acquire and process the sensors' data, and finally, show the obtained results.

This method includes different techniques for the accurately identification of ADLs. First, it makes use of the sensors in mobile devices to enable the capture of several physical and physiological parameters optimized with sensor data acquisition methods. Second, the data may be processed with lightweight methods implemented on a mobile application. However, some of the data captured are too complex to be processed locally, therefore they must be sent to a remote server, which carries out the data processing and returns the results back to the mobile device. However, these methods need a constant connection to the network. Besides, during the data acquisition and/or data processing stages, some failures may occur and data imputation and data cleaning techniques should be applied. Depending on the types of data captured, different data processing techniques must be applied. Currently, the case study in this PhD project is focused on the processing of audio for the identification of ADLs, applying audio fingerprinting methods. After that, data fusion techniques should be applied in order to merge all the features obtained by all the sensors available in the mobile device. After the data is fused and taking into account user's context, data mining, pattern recognition and/or machine learning techniques may be applied to create models for the identification of ADLs, which should be validated with the user's feedback.

3.2 Work in Progress

Until now, the state of the art about the roadmap of the proposed method has been already researched and published at [43], discovering the most and the best models used in other research studies.

At the same time of the capture of the data from the accelerometer, gyroscope, magnetometer, strength of the Wi-Fi signal and basic service set identification (BSSID) of the currently connected network, this project is aiming at creating a set of audio fingerprints for the identification of ADLs using the acoustic signal. The audio fingerprints will be based on the main features of the collected audio data using a mobile application specifically developed. This mobile application captures audio data, which is later stored in text files with signed or unsigned amplitude values, the accelerometer data, the values of the location based on a GPS receiver and the data related to the Wi-Fi connection, *e.g.,* the BSSID and the strength of the signal, which are also stored in text files for further analysis. At the end of this project, the data from all sensors are fused and the ADLs may be recognized with more accuracy.

The analysis of audio data for the identification of ADLs is now focused on four scenarios: watching TV, working at a software development company, sleeping and walking. We are collecting several samples for further analysis and creation of the metrics for audio fingerprints of each activity. In addition, the accelerometer, gyroscope, magnetometer, GPS, BSSID of Wi-Fi connected network and the strength of the Wi-Fi signal are captured for further analysis, in order to complement the recognition of the ADLs using the audio fingerprinting.

After the acquisition of a sample of audio data by the mobile device, it is processed to obtain the correspondent audio fingerprint that is then compared with the learned models based on the scenarios previously defined. However, each audio fingerprint obtained can match on several activities, thus distance metrics are applied. In addition, the use of multiple sensors for comparison are desirable, because it is expected to produce better results.

At the learning stage, the data should be captured by a smartphone and/or a smartwatch combined with the user's feedback and the users' agenda to create a set of signals' patterns and audio fingerprints with the characteristics of the data captured related to the different ADLs. After that, when a sample of the sensors' data is collected, it should be compared with the data stored in the collection, identifying the ADLs more accurately. Generally, the sensors' data related to the ADLs performed by a sedentary individual working at a software development company, includes the several ADLs, these are:

- between 0h00m and 8h00m, the activity is sleeping (during the working days);
- between 9h00m and 19h00m, the activity is working (during the working days, stopping the capture during the breaks);
- between 19h00m and 0h00m, the activity is watching TV (during the working days);
- between 15h00m and 19h00m, the activity is walking (during the weekend).

More samples should be acquired to compare and to create a large set of identifiable ADLs, in order to create the different methods to the development of a personal digital life coacher.

4 Discussion and Conclusions

The review of the state of the art related to part of the proposed method for this PhD project has been already published. It was focused on several concepts, such as the identification of ADLs, the existence of several mobile platforms, the diversity of the sensors available in off-the-shelf mobile devices, the concept of multi-sensor systems, the acquisition, processing, validation, cleaning, imputation, and fusion of the sensors' data, data analysis with pattern recognition techniques, and user's feedback.

The diversity of sensors available in mobile devices and the constraints related to the low resources of the mobile devices and the position of the sensors during the acquisition of the data are some challenges that this PhD project needs to address in order to create a robust method for the identification of the ADLs. The method that we expect to create during the project should be based on a subjective validation provided by the user combined with adjustable models in order to obtain better results, and statistically validated. Thus, several models should be implemented and tested for the different user's daily activities, such as: the k-NN for data imputation, and the Kalman filter for data fusion.

The research is now focused on the capture and process of audio data in order to create a set of consistent audio fingerprints for the identification of ADLs.

Acknowledgments This work was supported by FCT project UID/EEA/50008/2013 (*Este trabalho foi suportado pelo projecto FCT UID/EEA/50008/2013*).

The authors would also like to acknowledge the contribution of the COST Action IC1303 – AAPELE – Architectures, Algorithms and Protocols for Enhanced Living Environments.

References

1. Pires, I.M.S.: Aplicação móvel e plataforma Web para suporte à estimação do gasto energético em atividade física, in Informatics Department. University of Beira Interior, Covilhã (2012)
2. Pires, I.M., Garcia, N.M., Canavarro Teixeira, M.C.: Calculation of jump flight time using a mobile device. In: Proceedings of the HEALTHINF 2015 8th International Conference on Health Informatics, Lisbon, Portugal (2015)
3. Pires, I.M., et al.: Measurement of heel-rise test results using a mobile device. In: Proceedings of PhyCS 2015, 2nd International Conference on Physiological Computing Systems, Angers, France (2015)
4. Pires, I.M., Garcia, N.M., Flórez-Revuelta, F.: Multi-sensor data fusion techniques for the identification of activities of daily living using mobile devices. In: Proceedings of the ECMLPKDD 2015 Doctoral Consortium, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Porto, Portugal (2015)
5. Garcia, N.M.: A Roadmap to the Design of a Personal Digital Life Coach (2016)
6. Morgan, S.: Activities of daily living. In: Community Mental Health, pp. 141–158. Springer, Heidelberg (1993)
7. Kuspa, K., Pratkanis, T.: Classification of Mobile Device Accelerometer Data for Unique Activity Identification (2013)
8. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. ACM SIGKDD Explorations Newsletter **12**(2), 74 (2011)

9. Siirtola, P., Röning, J.: Recognizing Human Activities User-independently on Smartphones Based on Accelerometer Data. International Journal of Interactive Multimedia and Artificial Intelligence **1**(5), 38 (2012)
10. Dernbach, S., et al.: Simple and complex activity recognition through smart phones. In: 2012 8th International Conference on Intelligent Environments (IE). IEEE, Guanajuato (2012)
11. Salazar, L.H.A., et al.: A Systematic Literature Review on Usability Heuristics for Mobile Phones. International Journal of Mobile Human Computer Interaction **5**(2), 50–61 (2013)
12. White, R.M.: A Sensor Classification Scheme. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **34**(2), 124–126 (1987)
13. Xu, R., He, L.: GACEM: Genetic Algorithm Based Classifier Ensemble in a Multi-sensor System. Sensors **8**(10), 6203–6224 (2008)
14. Mo, L., et al.: Multi-Sensor Ensemble Classifier for Activity Recognition. Journal of Software Engineering and Applications **5**(12), 113–116 (2012)
15. Scalvini, S., et al.: Information and communication technology in chronic diseases: a patient's opportunity. Journal of Medicine and the Person **12**(3), 91–95 (2013)
16. Bersch, S.D., et al.: Sensor data acquisition and processing parameters for human activity classification. Sensors (Basel) **14**(3), 4239–4270 (2014)
17. Lim, L., Misra, A., Mo, T.: Adaptive data acquisition strategies for energy-efficient, smartphone-based, continuous processing of sensor streams. Distributed and Parallel Databases **31**(2), 321–351 (2012)
18. Dolui, K., Mukherjee, S., Datta, S.K.: Smart device sensing architectures and applications. In: 2013 International Computer Science and Engineering Conference (Icsec), pp. 91–96 (2013)
19. Sun, S., et al.: Literature review for data validation methods. Science and Technology **47**(2), 95–102 (2011)
20. Garza-Ulloa, J., Yu, H., Sarkodie-Gyan, T.: A mathematical model for the validation of the ground reaction force sensor in human gait analysis. Measurement **45**(4), 755–762 (2012)
21. Vateekul, P., Sarinnapakorn, K.: Tree-based approach to missing data imputation. In: 2009 IEEE International Conference on Data Mining Workshops, ICDMW 2009, pp. 70–75. IEEE, Miami
22. Patrician, P.A.: Multiple imputation for missing data. Res. Nurs. Health **25**(1), 76–84 (2002)
23. Ling, W., Dong-Mei, F.: Estimation of missing values using a weighted k-nearest neighbors algorithm. In: International Conference on Environmental Science and Information Application Technology, ESIAT 2009, pp. 660–663. IEEE, Wuhan
24. Jeffery, S.R., et al.: Declarative Support for Sensor Data Cleaning **3968**, 83–100 (2006)
25. Banos, O., et al.: On the use of sensor fusion to reduce the impact of rotational and additive noise in human activity recognition. Sensors (Basel) **12**(6), 8039–8054 (2012)
26. Pombo, N., et al.: Medical decision-making inspired from aerospace multisensor data fusion concepts. Inform. Health Soc. Care **40**(3), 185–197 (2015)
27. Tanveer, F., Waheed, O.T., ur Rehman, A.: Design and Development of a Sensor Fusion based Low Cost Attitude Estimator. Journal of Space Technology **1**(1), 45–50 (2011)

28. Bujari, A., Licar, B., Palazzi, C.E.: Movement pattern recognition through smartphone's accelerometer. In: 2012 IEEE Consumer Communications and Networking Conference (CCNC). IEEE, Las Vegas (2012)
29. Chikhaoui, B., Wang, S., Pigot, H.: A frequent pattern mining approach for ADLs recognition in smart environments. In: 2011 IEEE International Conference on Advanced Information Networking and Applications (AINA), IEEE, Biopolis (2011)
30. Kouris, I., Koutsouris, D.: A comparative study of pattern recognition classifiers to predict physical activities using smartphones and wearable body sensors. Technol. Health Care **20**(4), 263–275 (2012)
31. Zhu, C., Sheng, W.: Realtime recognition of complex human daily activities using human motion and location data. IEEE Trans. Biomed. Eng. **59**(9), 2422–2430 (2012)
32. Farrahi, K., Gatica-Perez, D.: Daily routine classification from mobile phone data. In: Machine Learning for Multimodal Interaction, pp. 173–184. Springer, Heidelberg (2008)
33. Hong, J.-H., et al.: An activity recognition system for ambient assisted living environments. In: Evaluating AAL Systems Through Competitive Benchmarking, pp. 148–158. Springer, Heidelberg (2013)
34. Phithakkitnukoon, S., et al.: Activity-aware map: identifying human daily activity pattern using mobile phone data. In: Human Behavior Understanding, pp. 14–25. Springer, Heidelberg (2010)
35. Nam, Y., Rho, S., Lee, C.: Physical activity recognition using multiple sensors embedded in a wearable device. ACM Transactions on Embedded Computing Systems **12**(2), 1–14 (2013)

Vox4Health: Preliminary Results of a Pilot Study for the Evaluation of a Mobile Voice Screening Application

Laura Verde, Giuseppe De Pietro and Giovanna Sannino

Abstract Mobile devices are rapidly becoming a part of everyday life as both communication and information tools. The use of m-health systems may be helpful for the collection of a set of data and information that can enable the individual to be proactive in the management of his/her own health, focusing on wellness and preventive behaviors, and so improving quality of life. We have realized an app, called Vox4Health, able to perform a fast, portable and simple voice screening test, due to the rapid increase in the incidence of voice disorders. The voice screening test is especially useful for teachers, singers, actors and other professional voice users who use their voice in their own activities, this vocal abuse possibly increasing the risk of suffering from dysphonia. Vox4Health analyzes vocal signals in real time, allowing users to self monitor the state of their own vocal health, thanks to an appropriate methodology. In this paper we present a pilot study conducted in two regions of Italy. The aim of the study is to evaluate the accuracy, in terms of classification capability, the usability and the degree of satisfaction with the developed app. The preliminary results show an average score of 81.3 for usability and average values ranging from 1.48 for dependability to 2.12 for perspicuity calculated in accordance with the User Experience Questionnaire.

Keywords mHealth application · Fundamental frequency estimation · Pilot study · Usability · Voice screening test

L. Verde
Department of Technology, University of Naples Parthenope, 80143 Naples, Italy
e-mail: laura.verde@uniparthenope.it

G. De Pietro · G. Sannino(✉)
Institute of High Performance Computing and Networking (ICAR),
National Research Council of Italy (CNR), 80131 Naples, Italy
e-mail: {giuseppe.depietro,giovanna.sannino}@na.icar.cnr.it

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_15

1 Background

Dysphonia is a disorder that alters the voice quality. The symptoms of this disorder can be characterized by objective signs, either acoustic (alterations in intensity, frequency, timbre or texture), clinical (morphological or dynamic), physical (pharyngeal-laryngeal paresthesia), or psychological (a feeling of unpleasantness or inadequacy of one's own voice), occasionally or constantly present, in all or in particular communicative situations. The incidence of voice disorders is increasing rapidly, such problems now affecting nearly one-third of the population [1]. Professional voice users (singers, actors or teachers) represent the most affected category [2]. Unfortunately, too few people know about dysphonia and the causes of this disorder, and for these reasons only a few sufferers consult a speech specialist.

Speech therapists evaluate the presence of dysphonia by performing several tests including an acoustic analysis, in accordance with the SIFEL protocol [3] released by the Italian Society of Logopedics and Phoniatrics, agree with the instructions of the Committee for Phoniatrics of the European Society of Laryngology. The acoustic analysis is useful to quantify the voice features through estimations of specific parameters like the Fundamental Frequency (F_0), or jitter and shimmer. The parameters are calculated from a recording of the vowel 'a' of five seconds in length. Each of these parameters represents a particular feature of the human voice. For example, the F_0 indicates the oscillations of the vocal folds, and its alteration possibly indicating the presence of voice disorders.

We have realized a mobile health (m-Health) application [4], that we have called "Vox4Health", that is able to estimate the F_0 of the voice signal and to classify it as healthy or pathological through a methodology, opportunely implemented for a smartphone or a tablet. This would be a simple and fast instrument to perform a real-time voice screening able to support people for the self-monitoring of the state of their voice, and to suggest healthy behaviors to prevent or reduce voice health problems. The objective of the study detailed in this paper is to test Vox4Health, evaluating the performance of the embedded methodology for the F_0 estimation, the usability, and the user's satisfaction with the realized app. We have drawn up an appropriate trial protocol, a detailed document in which we have described the complete execution plan, focusing on organizational and ethical aspects. Based on this analysis of the literature, we have built a trial protocol in accordance with the SPIRIT 2013 Statement [5], which provides some guidelines on the performance of a clinical study.

2 Vox4Health

The system realized is able to estimate the F_0 from the speech signal and to classify the analyzed voice as healthy or pathological. In the following subsections, descriptions of the F_0 estimation and the functionalities provided by the app are presented.

2.1 Fundamental Frequency Estimation

The developed methodology to estimate the F_0 from vocal signal consists of several operations, such as the application of a difference function and a cumulative mean normalized difference function, the calculating of some threshold values, the application of a parabolic interpolation, and the calculation of an average value to estimate the F_0 parameter. This methodology is based on the Yin method [6] and is detailed in [4] and in [7]. In this latter manuscript, we have described the improvements in the accuracy of the methodology achieved by introducing a Finite Impulse Response (FIR) filter to remove the noise components that can compromise the useful voice signal.

2.2 Functionalities of the App

The methodology is included in a mobile application, called Vox4Health, developed by using the Java Programming Language. To access such functionalities, the user must be registered, through the insertion of his/her information such as name, surname, e-mail address, password, date of birth and gender. These data, in particular the e-mail address and the password, are necessary to log-in to the Wellness Server. The connection to the Wellness Server can store a knowledge base consisting of data and documents that are shared with the other apps developed within the project Smart Health 2.0 [8], to guarantee an environment of cooperation and information sharing with the unique common goal of the improvement of the well-being of the patient. When the user accesses the app, he/she can perform various procedures, such as visualizing the General Indications, acquiring an Audio file, loading an audio, compiling some questionnaires required by the SIFEL protocol, or loading a report. Interested readers can find detailed information about the functionalities in [4].

3 The Trial Protocol

The developed protocol consists of several sections:

- **Administrative**: In this first section, all administrative information are reported, such as the descriptive title that identifies the study design, useful for the trial identification, the sources of financial support, and the roles and responsibilities of the contributors to the process.
- **Introduction**: Here a description of the research question, including a the set of the relevant published studies and the outcomes of the process, is reported. The outcomes are fundamental for the trial process and the interpretation of results. The main outcomes are indicated as primary outcomes, while the remaining outcomes are described as secondary outcomes.

– **Methods**: In this section the procedures to be followed during the process, identifying the location of the study (identification and description of the operating structures), the eligibility criteria for participants (inclusion and exclusion criteria useful to identify the study population), the timing, the recruitment and allocation of participants, the plans for assessment and the collection of data are specified.
– **Ethics and Dissemination**: The protocol also provides for the definition of plans for the ethical approval of the trial process and possible changes to the protocol in the Ethics and Dissemination section. Moreover, the rules and behavioral norms of the participants are specified, approved by the signing of the informed consent.
– **Appendix**: Finally, in the Appendix the necessary documents for the trial process, such as the informed consent, the information sheet, the test for usability and the anamnestic form, are included.

4 Pilot Study for Vox4Health

The trial was conducted in certain urban areas in the Campania and Calabria regions, in detail in Naples, Cosenza, Crotone, San Giovanni in Fiore, and Lamezia Terme. We divided the study into five phases:

1. Participants Recruitment
2. Identification of Eligible Participants
3. Participants Registration
4. 1st Meeting: Installation of the app
5. 2nd Meeting
6. 3rd Meeting: Data collection

We started with the participants' recruitment phase by means of the promotion of the testing program through information campaigns as documented in [9, 10, 11], and through the use of tools such as posters, news published on the project web-site [8], brochures, and news published on the main social networks, such as Linked-In and Facebook. For each person who expressed an interest in participating in the trail, possession of the eligible criteria requirements was verified. If this was confirmed, the individual was able to access the second phase of the trial. Here, we provided information on the scope of the study and the procedures of the trial process and informed the participants that it was possible to interrupt the experimental practice whenever they wanted. The interested individuals were then registered for the trial by signing an anamnestic form with all their personal, medical and contact information, a form for confirmation of the informed consent, and the privacy statement essential for guaranteeing data protection. The anamnestic form is necessary in order to detect any similarities between the individual's distinctive characteristics and the onset of the voice disorders. The next steps of the trial process were three meetings with the participants distributed over two months. During the first meeting we installed the app on the mobile devices and explained the functionalities of Vox4Health. It is important to note that the Android operating system is required for the installation and use of the app. For each participant who did not have an Android-based smartphone, we provided a Mediacomm Phonepad G501 smartphone on loan for use throughout the duration of the study. In the second meeting, in order to involve the participants in the trial process, they were updated on the progress of the project, as well as provided with a new improved version of Vox4Health realized in accordance with the indications suggested by the participants during the previous meeting. In the third and last meeting we collected all the data from the mobile devices of the participants, and additionally we talked with them encouraging them to discuss the use of the application, any difficulties encountered during the trial, the perceived quality of the information provided, and any preferences for different features. In this latter meeting we also distributed the usability test questionnaires, asking the participants to answer in a very spontaneous way. All the questionnaires were anonymous.

4.1 Methods

We recruited 49 volunteers into the study. In particular, 26 people were involved in Naples (16 females and 10 males) and 23 (20 females and 3 males) in the four districts of the Calabria region.

The experimental tests were performed individually. After the installation of Vox4Health on the mobile devices each participant made a short recording on the Wellness Server. This registration is required to access the functionalities of the app. Each participant was given a short period of familiarization with the app, about ten minutes, during which he/she was able to examine the offered features. Successively, we asked the participants to carry out a series of application-specific tasks that included the following:

1. Access to the app;
2. Estimation of the Fundamental Frequency by the recording of a vocalization of the vowel “a” five seconds in length without any interruption of sound;
3. Completion of the Voice Handicap Index (VHI) questionnaire [12] to estimate the self-perception of their own voice;
4. Completion of the Reflux Symptom Index (RSI) questionnaire [13] to estimate the self-perception of the presence and severity of extra-esophageal reflux;
5. Loading of a saved audio capture;
6. Loading of a saved report.

All analyses, including the F_0 estimation and the calculation of the indexes (VHI or RSI), are performed in real time. All evaluations are shown to the user providing a feedback about the analysis performed, summarized in a pdf report file, suggesting the corrective actions capable of improving the state of well-being. Concerning the completion of the two questionnaires, the VHI and RSI, each participant must answer the questions submitted, to which a specific score depending on the severity of the symptom corresponds [12] [13]. During each session, the users were always encouraged to discuss the ease of use of the application, the difficulties encountered,

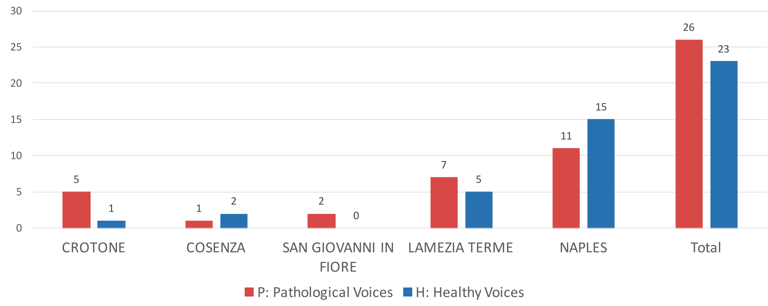


Fig. 1. Classification Results and Questionnaires Results.

the perceived quality of the information provided, and any preferences for different features. These assessments were collected by means of an evaluation of the post-usability test questionnaires, described in the following subsection. All data, including audio recordings, field notes, usability questionnaires with participants' comments, were saved in anonymous mode, and filed and elaborated to evaluate the results obtained. Each participant was uniquely identified with a code.

4.2 Voice Classification Analysis

To classify a voice as healthy or pathological, we used a range of values selected in collaboration with a group of experts from the Department of Otorhinolaryngology at the University Magna Graecia of Catanzaro (Italy) who are involved in the project. In detail, the healthy range is 200-250 Hz for female voices and 100-150 Hz for male ones. All values of F_0 calculated by the app outside of this range are considered pathological voices. The analysis of the participants voice signals indicated that, out of 49 analyzed voices, 23 were considered to be healthy and 26 pathological. The classification results are summarized in figure 1. There were 46 completed self-assesment questionnaires. From an analysis of the VHI questionnaires it was ascertained that 3 participants considered themselves to have a disorder with their voices, that impacts on the quality of their lives. From an analysis of the RSI results, it was observed that 11 people judged themselves to suffer from extra-esophageal reflux.

4.3 Usability Evaluation of the App

A well-designed medical mobile application is generally useful and usable [14]. Usability is usually measured in terms of three attributes: effectiveness, efficiency and satisfaction. The first, effectiveness, indicates the accuracy and completeness with which users can achieve specified goals. Efficiency, instead, can be evaluated by estimating the resources expended in relation to the accuracy and completeness with which the users achieve these goals. Finally, satisfaction refers to the user's subjective assessment of satisfaction. To evaluate these features we used two usability test questionnaires:

– The **System Usability Scale (SUS)** [15] was used to assess the usability of the Vox4Health application. It consists of a simple and reliable tool, a Likert scale, which provides a comprehensive understanding of the subjective assessments of usability, assigning a score to each question indicating the degree of agreement or disagreement on a 5-point scale, providing an overall score between 0 and 100. The values equal to 1 represent a level of strong disagreement with the assessment proposed, while the values of 5 represent strong agreement.
– The **User Experience Questionnaire (UEQ)** [16], employed to assess, in particular, six basic aspects such as the Attractiveness of the product (a general impression about the product, if for example users like or dislike it), Perspicuity (if for example it is easy to learn, and easy to understand), Dependability (if the user controls the interaction, if the interaction with the product is secure and predicable), Efficiency (is it fast, and organized?), Novelty (is it creative and innovative?) and Stimulation (is it interesting?). For each of these parameters there is associated a score, ranging between -3 and +3. -3 represents the most negative response, 0 a neutral response, and +3 the most positive response. Scale values above +1 indicate a positive impression of the users, while values below -1 represent a negative impression. Since, in the practice setting, extreme values are rarely observed, the examined scale means are in general in the range from -2 to +2. The scores between -0.8 and +0.8 represent a neutral evaluation, values greater than 0.8 represent a positive rating while those below 0.8 a negative assessment.

In Figure 2 the average given score for each question of the SUS questionnaire is reported. All items are rated on a 1-5 Likert scale, where 1 corresponds to the most negative answer (not enjoyable/ not helpful/ not liked), while 5 represents the most positive answer (enjoyable/ helpful/ liked). The analysis of the SUS tests shows that the application was well-received by the participants achieving a SUS score average of about 81.3. The app was considered easy to use, well organized and with a clear design, as shown in Figure 4, where the questions relating to the ease of use and the integration of the functionalities of the app achieved high scores (respectively 4.50 and 4.32). All the participants in the study used the application for about two months. After this experience, they completed the usability questionnaires, both the SUS and UEQ.

The maximum running speed of the execution of each test is approximately 70 ms and this was considered adequate, as well as the navigation being clear and unambiguous. These results were corroborated by the evaluations of the UEQ questionnaires, which confirmed that the app achieved good results for the six analyzed variables, the degree of interest in the product, perspicuity, dependability, efficiency, novelty and stimulation, as shown in Figure 3. In this figure we can note that perspicuity achieved a good score (about 2.12) which means that for the participants it was easy

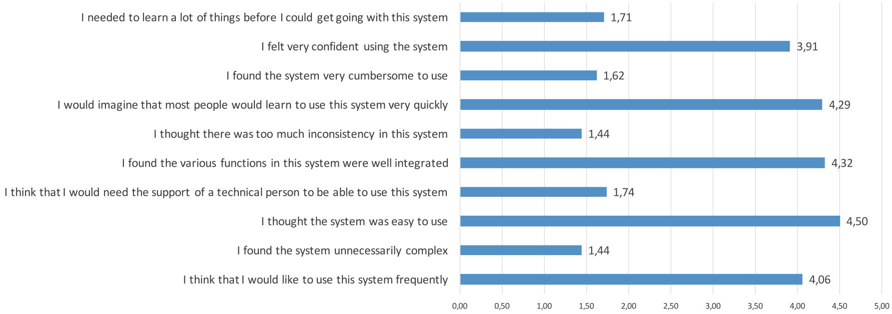


Fig. 2. The average given score for System Usability Scale (SUS) questions

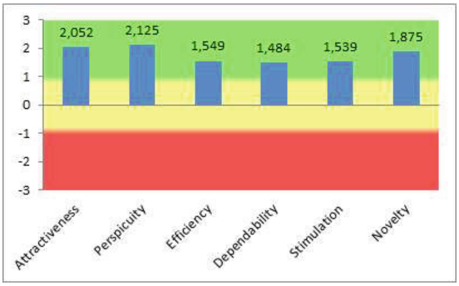


Fig. 3. User Experience Questionnaire (UEQ) results

to understand how to use the app and to become familiar with its features. Furthermore, the app was liked, interesting and appealing (the scores for attractiveness and stimulation were, respectively, 2.05 and 1.53).

5 Conclusions

Dysphonia is a voice disorder that can impact on the social and professional life of sufferers people. The evaluation of the vocal health is performed thanks to the acoustic analysis, an important examination provided in accordance with the SIFEL protocol. It consists of the quantitative evaluation of characteristics parameters such as the Fundamental Frequency. We have realized a methodology able to estimate F_0 and to discriminate between healthy voices from pathological ones, and we have embedded it in a mobile application called Vox4Health, useful to easily perform a screening test.

In this paper we have presented a trial study conducted to test Vox4Health among a significant number of volunteers. The experimental phase has been useful to evaluate the usability of the app and the users' level of satisfaction,calcutated by using post-usability test questionnaires. Additionally, this study will allow us to evaluate also the reliability of the app in connection with the estimation of the F_0 and the classification of the voice between healthy and pathological. This will be undertaken in the next feature, thanks to the University "Magna Graecia" of Catanzaro and particularly their speech therapist staff who will perform a specialized medical examination of each participant in this study. In this way, we will be able to calculate classification performance parameters like accuracy, sensitivity and specificity. The results will be compared with the most well known tools for the estimation of F_0 exiting in literature.

The results show that Vox4Health is considered easy to use, it has a clear and uncluttered screen design, and incorporates typographic elements that communicate meaning and help users understand context and interactivity. Both questionnaire results are good and all participants at the trial have expressed interest in the use of the app and have enjoyed the intuitiveness and the ease of use of all its functionalities.

Acknowledgement The authors would like to acknowledge the project "Smart Health 2.0" PON04A2_C for its support of this work. Additionally, the authors wish to thank Prof. Nicola Lombardo, Department of Otolaryngology-Head and Neck Surgery of the University "Magna Graecia" of Catanzaro (Italy) involved in the SmartHealth 2.0 project.

References

1. Roy, N., Merrill, R.M., Gray, S.D., Smith, E.M.: Voice disorders in the general population: prevalence, risk factors, and occupational impact. The Laryngoscope **115**(11), 1988–1995 (2005)
2. Van Houtte, E., Claeys, S., Wuyts, F., Van Lierde, K.: The impact of voice disorders among teachers: vocal complaints, treatment-seeking behavior, knowledge of vocal care, and voice-related absenteeism. Journal of Voice **25**(5), 570–575 (2011)
3. Ricci Maccarini, A., Lucchini, E.: La valutazione soggettiva ed oggettiva della disfonia: il protocollo sifel. In: Relazione ufficiale al XXXVI Congresso Nazionale della Società Italiana di Foniatria e Logopedia (2002)
4. Verde, L., De Pietro, G., Veltri, P., Sannino, G.: An m-health system for the estimation of voice disorders. In: 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE (2015)
5. Chan, A.-W., et al.: SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. Bmj **346**(2013), e7586 (2013)
6. De Cheveign, A., Kawahara, H.: YIN, a fundamental frequency estimator for speech and music. The Journal of the Acoustical Society of America **111**(4), 1917–1930 (2002)
7. Verde, L., De Pietro, G., Veltri, P., Sannino, G.: A noise-aware methodology for a mobile voice screening application. In: 2015 6th IEEE International Workshop on Advances in Sensors and Interfaces (IWASI), pp. 193–198. IEEE, June 2015
8. Smart Health 2.0, SH.2.0. (2014). http://www.smarthealth2-0.com
9. http://www.cn24tv.it/news/116057/controllare-il-proprio-stato-di-salute-con-un-app-nasce-health-2-0.html
10. http://www.crotoneinforma.it/pgn/news.php?id=6165#.Vd72U_ntmko
11. http://www.comune.cosenza.gov.it/archivio10_notizie-e-comunicati_0_15502.html
12. Forti, S., Amico, M., Zambarbieri, A., Ciabatta, A., Assi, C., Pignataro, L., Cantarella, G.: Validation of the Italian Voice Handicap Index-10. Journal of Voice **28**(2), 263–e17 (2014)

13. Belafsky, P.C., Postma, G.N., Koufman, J.A.: Validity and reliability of the reflux symptom index (RSI). Journal of Voice **16**(2), 274–277 (2002)
14. Patricia Arellano, J.B., Elias, B., Houser, S., Martin, T., Head, H.: Selecting a Mobile App: Evaluating the Usability of Medical Applications. Presented at the mHIMSS App Usability Work Group (2012)
15. Brooke, J.: SUS-A quick and dirty usability scale. Usability Evaluation in Industry **189**, 4–7 (1996)
16. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. Springer (2008)

Assessing Interpersonal Trust in an Ambient Intelligence Negotiation System

Marco Gomes, Cesar Analide and Paulo Novais

Abstract This paper describes an approach to assess and measure trust based on a specific Ambient Intelligence environment. The primary aim of this work is to address and expand on this line of research by investigating the possibility of measuring trust based on quantifiable behavior. To do so, we present a brief review of the existing definitions of trust and define trust in the context of an Ambient Intelligence (AmI) scenario. Further, we propose a formal definition so that the analysis of trust in this kind of scenarios can be developed. Thus, it is suggested the use of Ambient Intelligence techniques that use a trust data model to collect and evaluate relevant information based on the assumption that observable trust between two entities (parties) results in certain typical behaviors. This will establish the foundation for the prediction of such aspects based on the analysis of people's interaction with technological environments, providing new potentially interesting trust assessment tools.

Keywords Ambient Intelligence · Negotiation · Trust

1 Introduction

Trust is understood as a complex phenomenon and it is widely accepted as playing a significant role in human social relationships. Analyzing the phenomenon and its formal definitions we can find a profusion of interpretations containing different trust types or facets, with different properties. A situation in which requires different models for analysis. This statement is particularly evident when one go deeper into the primary disciplines concerned with trust relationships (psychology, sociology, etc.). This abundance of meanings inevitably leads to a degree of uncertainty about what is meant by trust, creating an enormous conceptual and terminological confusion.

M. Gomes(✉) · C. Analide · P. Novais
Department of Informatics, University of Minho, Braga, Portugal
e-mail: {marcogomes,cesar.analide,pjon}@di.uminho.pt

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_16

In fact, the most consistent thing about trust seems to be when it is absent [1]. Bearing in mind these diverging views as to what constitutes the appropriate definition of trust, the present work focuses on some specific definitions of trust, which are simplified so that these views can be captured algorithmically. Meanwhile, while a variety of definitions of the term *trust* have been suggested, in this work we will use a specific definition in which trust is seen as a rich and a complex mental attitude of *x* towards *y* as for a given action and goal [2]. Accordingly to Castelfranchi [2], this approach consists of evaluations of *y* and the situation, and of expectations about *y*'s mind, behaviour and possible results. This includes, of course, the assumption that trust should be understood and can be interpreted by framing its natural subjectivity and the information needed at a particular time and for a specific context in a computer-based model: an abstraction that have the *power* to represent data in terms of entities and relationships relevant to a domain of inquiry (trust).

However, it is needed to find a way of evaluating (measure) trust to guide the research effort in the correct direction. Facing this issue, a critical question must be raised: how to balance the subjective nature of trust with the objectivity-dependent nature of a computer system? In other words, how can a computer system deal intelligently with this kind of subjectivity? Well, some glimpses of the ways that computer systems meet subjectivity can be found in the "expert" literature. This is exemplified in the work undertaken by Rosalind Picard in which she stressed that this endeavour is challenging but achievable [3]. Afterwards, she outlined a strategy for computer systems to cope with subjectivity issues. Specifically, it is proposed a three-fold approach: that they [computer systems] will need to (1) share some of the common sense of the user, (2) observe and model the user's actions, and (3) learn from these interactions. To summarize, this approach suggests that subjectivity is expressed as the user interacts with the system during a succession of queries. These involving inputs of the user can be tracked, modelled, and used to retrieve data consistent with changing requests. In our opinion, this strategy seems to be suitable to be applied so we planned to follow it in our work in ways to overcome the further issue. Hence, our motivation to examine trust is two-fold. First, the present study aims to address and expand on this line of research by investigating the possibility of measuring trust based on quantifiable behavior. To do so, we present a brief review of the existing definitions of trust and define trust in the context of an Ambient Intelligence (AmI) scenario. Further, we propose a formal definition so that the analysis of trust in this kind of scenarios can be developed. Thus, it is suggested the use of Ambient Intelligence techniques that use a trust data model to collect and evaluate relevant information based on the assumption that observable trust between two entities (parties) results in certain typical behaviors.

2 Measuring Trust using an Algorithmic Approach

As suggested in [4], trust literature can be categorized based on three criteria: (i) trust information collection, (ii) trust value assessment, and (iii) trust value dissemination. Each, in turn, can be further classified: trust information gathering from three sources, namely (i) attitudes, (ii) behaviors, and (iii) experiences; trust value assessment according to the data model, namely (i) graph, (ii) interaction, and (iii) hybrid; and trust value dissemination into trust-based recommendation and visualization models. About this work, and taking into account the Sherchan's strategy [4], we will highlight the trust definitions, trust types, properties, and measurement models from the perspective of the Computer Science and Economics disciplines. Such approach seeks to focus on some particular aspects of trust measurement, which are simplified so that these properties can be captured algorithmically. The simplifying process is attached to the need for obtaining and/or quantifies trust using detecting statistically significant trust-like behaviors. Therefore, the basis for our study is a proposition that trust results in characteristic interaction behavior patterns that are statistically different from random interaction in a computerized conflict and negotiation management system [5]. However, regarding the relationship between trust and negotiation one must stress that different research streams view trust in different ways depending upon the relationship under consideration [6]. So another important decision was to reduce the study of *trust domain* to the study of *interpersonal trust*, transforming the challenge of measuring trust slightly more accessible.

Facing these challenges, how trust measurements can be classified? Accordingly to [4], in general, it can be classified into two broad categories: "user" and "system". For the scope of this present work, we will consider only the notion of "user" trust is derived from Psychology and Sociology, with a standard definition as "a subjective expectation an entity has about another's future behavior" [4]. This implies that trust is inherently personalized. In this sense, trust is *relational*. As two individuals interact with each other frequently, their relationship strengthens, and trust evolves based on their experience. Following Adali proposition [7] related to interpersonal trust, the main type or facet of trust understudy will be the basis for our proposal. Hence, the focus will be on the following proposition: is possible to observe that trust between two entities A and B will result in certain typical behaviours that can be statistically captured. In other words, our aim is to quantitatively measure dyadic trust (trust between two entities) based on observed behaviour in a negotiation process. These behaviours are not only an expression of trust but can also facilitate the development of further trust. The simplest such behaviour is just interaction, in which an action occurs as two or more entities have some kind two-way effect upon one another. Regarding trust evaluation models, our approach to trust computation will be based on an interaction-based trust model. In this case, our interaction-based model evaluates trust based on the interactions performed in a computer-based conflict management system. Then trust-related information will be captured and assessed following the strategy aforementioned. The sources and the means to assess trust can be resumed to:

- **Source of trust information gathering**: Behaviours. Why? Because user behaviours are an important aspect of trust. Another reason is as they are identified by patterns of interactions they can be easily captured by an algorithm running on a computer system. Therefore, this will be our main source of information gathering. For example, in a negotiation scenario if a party is an extremely active participant and suddenly stops participating, this change in behaviour (he interaction is interrupted) is noticeable and might imply that this party's trust in the other party or with the party with whom he/she had been frequently interacting with has decreased.

- **Trust value assessment**: Regarding the techniques used to measure (compute) trust, they can be broadly classified into statistical and machine learning techniques, heuristics-based techniques, and behaviour-based techniques. In the present work, we will use behaviour-based techniques to assess trust. The reason for this choice lies in the fact that our measure of trust is based on quantifiable behaviour. So it seems obvious to choose assessment techniques based on this proposition.

After we have defined which directions and techniques we will embrace to measure trust, let us then formally define our proposal based on Adali's approach. This formalization is applied to the context of a negotiation, a process for two (or more) parties to find an acceptable solution to a conflict. Within a negotiation process, each party can make several proposals and exchange an unlimited number of messages. For the context of the interaction in a negotiation, the input is the proposal stream in a negotiation process, specified by a set of *4-tuples*,

$$\langle sender, receiver, message, time \rangle \tag{1}$$

note that we pretend to study the problem of trust purely from the observed interaction statistics, using no semantic information. Meanwhile, in our formalization, in an interaction between parties some semantic aspects (e.g. message) are considered in order to provide posterior semantic analysis. The output considered here is a set T induce from these inputs. The participants of the negotiation are represented by the elements of this set.

Despite what common sense stands in a *regular* situation (which more often interactions occur, the more likely that a trust relationship is likely to exist or to develop), the presence of distrust can imply a lack of interaction in conflict mitigation process. Assuming that, we postulate in this present work that the shorter and less balanced (means the average number of times that two entities interacts within the process) a interaction is between two parties, the more likely it is that they have a trust relationship; in addition, the more interaction there are between such a pair of elements, the less tightly connected they are. The fundamental task is first to identify when two elements of T set are interacting. Let A and B be a pair of users, and let $P = \{t_1, t_2, ..., t_k\}$ be a sorted list of the times when a message was exchanged between A and B. Therefore the average time between messages is defined as $\tau = (t_k - t_1)/k$.

The measure of trust will be based on the interactions in I, obeying the following postulates:(1)Longer interactions imply less trust;(2)More interactions imply less trust; and (3)Balanced participation by A and B implies less trust. We define the relational trust $R_I(A, B)$ as follows:

$$R_I(A, B) = \sum_{i=1}^{l} \|I_i\| \cdot H(I_i) \tag{2}$$

Where $H(I_i)$ is a measure of the balance in the i interaction contained in I. We use the entropy (measure of the amount of information that is missing in the flow of interaction) function to measure balance:

$$H(I_i) = -p \log p - (1 - p) \log(1 - p), \qquad (3)$$

where $p(Ii)$ is the fraction of messages in the interaction Ii that were performed by A. The complexity of the algorithms for computing relational trust is $\Theta(|D| \log |D|)$, where $|D|$ is the size of the interaction stream.

3 Ambient Intelligence Applied to Trust Assessment

Ambient Intelligence has many uses in a wide domain. Under this paradigm, computational power is seamlessly embedded into the environment, ultimately creating computational environments that implement their life-cycle in an ideally invisible way for the user. Our primary aim is to develop AmI platform that will support the already electronic negotiation systems by providing relevant context information derived from the environment. In that sense, this work was decided to develop an environment that could be sensitive and responsive with both parties of a negotiation platform. As a result, this environment allows that all the components can be combined to implement complex functions. Namely, monitoring the negotiation process and perceiving how each issue or event is affecting each party, namely knowing the interpersonal trust of each party in each round.

Towards an intelligent conflict support system an ambient intelligence system was developed. The general working is to sense conflict context, acquire it and then make reasoning on the acquired context and thus acting in on the parties' behalf. To achieve this, the system builds up a profile and can link that profile subsequently with the correct individual performance within the conflict process that is monitored by the system. In other words, while the user conscientiously interacts with the system and takes his/her decisions and actions, a parallel and transparent process takes place in which contextual and behavioural information is sent in a synchronized way to the conflict support platform. The platform after converting the sensory information into useful data allows a contextualized analysis of the user's data. The contextualized analysis of user's data is critical when the data is from heterogeneous sources of diverse nature like sensors, user profile, and social media and also at different timestamps. To overcome some of these problems, the features are extracted from multiple sensor observations and combined into a single concatenated feature vector that is introduced into different classification modules (conflict styles, trust analysis, etc.). The multimodal evidence are integrated using a decision level strategy. Examples of decision level fusion methods employed in this work include weighted decision methods and machine-learning techniques and are detailed in previous work [8].

3.1 Case-Study: A Negotiation Scenario

A negotiation *scenario* should specify everything we know about the problem being addressed to frame the interactions that occur within. Taking this into account, a technological framework aimed to support the decision-making, by facilitating access to information such as the negotiation style of the parties or their social context, was adapted. In this work, it is introduced a new module that takes into account the context using trust analysis. Moreover, at this point, is highlighted that the primary objective of this research work is to identify and measure the users' interpersonal trust, to correlate to their negotiation performance and how it can be pointed out. Therefore, an experiment was set up in which we tried to estimate all the relevant aspects of the interaction between the individual that occur in a sensory rich environment (where contextual modalities were monitored). The participants of the proposed experiment were volunteers socially connected with our lab members. Twenty individuals participated, both female and male, aged between 19 and 42. The first step of the experiment was to ask the volunteers to fill in a small individual questionnaire. The next step was the monitoring of the individuals' interaction with the developed web-based negotiation game (in which subjects perform two distinctly different roles). During the experiments, the information about the user's context and performance (extracted using models based on behavioral and contextual monitoring [9], [10]) was provided through a monitoring framework, which is customized to collect and treat the interaction data. The participants played the web-based game through computers that allowed the analysis of the described features.

3.2 Results

After the experiment has been performed, the first step was to run some analysis to compare two sets of data under study: the data collected through the application of the questionnaire (measuring the participant's relationships) and the data gathered through the web-based negotiation game. According to our postulates, the basis for this analysis was the assumption that if A and B have a strong relationship then $R_I(A, B)$ value is below the median of the calculated trust for all the pairs within the interactions data set. In other words, if one pair of participants has a high degree of relationship then the same pair of participants, during the game, will perform fewer interactions than the median of total interactions per game. At this point, it should be highlighted that to apply non-parametric statistical analysis the raw data were pre-processed and it was subjected to tests. Therefore, the outcomes were compared using the Mann-Whitney U test (compares the central tendencies of two independent samples), given the fact that most of the distributions are not normal. The null hypothesis is thus: H_0 = The medians of the two distributions are different. For each two distributions compared, the test returns a p-value, with a small p-value suggesting that it is unlikely that H_0 is true. For each parameter (pair of participants), data from both samples is compared. In all the tests, a value of $\alpha = 0.05$ is used. Thus, for every Mann-Whitney test whose $p-$value $< \alpha$, the difference is considered to be statistically insignificant, i.e., H_0 is rejected. Consequently, the results have shown that no (statistically) important difference between data from the two samples were found. In other words, it means that our assumptions were valid. Another aim of this present work was to study the link in the relational trust between opponents and their behaviour exhibited in a negotiation scenario. So to analyse if trust relationships influence the negotiation performance, in the preliminary data analysis, the experimental data is organized into two groups based on the analysis of the trust measurements. One group contains the collection of some experimental data about how a user (A) behaves when he/she negotiates with someone (B) in which $R_I(A, B)$ have low value ($R_I(A, B) <$ median). In that sense, this approach will enable the establishment of a baseline for comparison with the second group, which comprises the data gathered from parties that negotiate with someone that haves high R_I values ($R_I \geq$ median). In particular, to statistically deal with data concerning the utility values of the parties' proposals, it was necessary to convert to an arbitrary numeric scale (0 is the least favorable style for the resolution and four the most favorable style). This type of scale means that the exact numeric quantity of a particular value has no significance beyond its ability to establish a ranking over a set of data points. Therefore, it was built rank-ordering (which describes order), but not relative size or degree of difference between the items measured. This was a mandatory step to make the data suitable for statistical and machine-learning techniques.

Moreover, the analysis shows that there is an apparent difference between the two groups regarding the negotiation styles exhibited during the game. One conclusion is that when participants share a significant trust relationship (low R_I value) the frequency of collaborative behaviours is far superior (49%) than otherwise (24%). In a similar analysis, but now concerning the roles played by participants, we conclude that the sellers are much more competitive than buyers (57% vs. 29%) while buyers are primarily collaborative. To interpret the significance of these results it is important to recall that participants were asked to negotiate a favorable deal in a competitive and win-lose scenario. Nevertheless, it is shown that when participants have a significant trust relationship they are more likely to transform it into a win/win situation. Something visible in the final results of the negotiations. On the one hand, we find that 100% of the agreements made by parties with a relevant trust relationship accomplished a successful deal, i.e., between the range of solutions that would benefit both. On the other hand, only 50% of negotiations that occurred between untrusted opponents (high R_I value) reached a mutual benefits agreement. It may be that they assumed they had to negotiate and get the best price (win/loose). But that was not the objective. Their objective was to negotiate a deal so they would not go bankrupt (win/win).

The preliminary evidence suggests a basis for expecting a connection between trust relationship and the use of negotiation styles. Despite these results, we still do not know much about how this kind of influence might facilitate (or inhibit) positive negotiation outcomes. Therefore, we will perform more and deeper experiments to understand how to collect and analysis relational ties that can influence negotiation performance.

4 Conclusion

We aim at firstly to identify and apply an algorithm for measuring interpersonal trust; secondly, to validate this approach opposing data collected from a questionnaire with data gathered from a web-based negotiation game to statistically study the correlation between mutual trust and conflict styles. From the experiment outcomes, the findings highlight the potentially quantifiable measurements of trust to further the understanding of negotiation dynamics. They pointed out relationships between the features being monitored and the participants' relationships elicited through a small questionnaire. These findings have the potential to enable the characterization of individuals and enhance negotiation performance. Thus the identification of trust relations between opponents in negotiation scenario in which can influence the negotiation performance is the main contribution of this work. Meanwhile, we can also conclude that these results are preliminary in the sense that there is more information that one can retrieve from the collected data namely through a more profound semantic analysis. This type of analysis could considerably enhance the trust measurements. Furthermore, due to the small sample size used in the current study, some caution must be taken when interpreting the results of the statistical analysis presented and underpinning the conclusions. In that sense, additional limitations of the current research must be pointed out. First, the participants were recruited from a particular population (that are socially related to our lab members)- a population that may limit the generalizability of the results. Admittedly, the participants of the experiment may not be representative of negotiation parties in general. Consequently, we are unable to demonstrate the causality of the variables conclusively. Moreover, it is possible that individual differences (i.e., personalities) might have influenced the impact of the results. Second, we tested all of the variables at the individual level that limited us from conducting global level analysis, which could provide us more variance of the data. This can be seen as another drawback of our study. Also, the data were collected through self-reported surveys at one time, which is subject to common method variance problem. Finally, the computational facet of this work should provide an understanding of the difficulties in algorithmically capturing and computing interpersonal trust in an AmI environment. A more comprehensive and in-depth study to provide theoretical advances, as well as implement technological solutions, is yet under development.

Acknowledgements This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT - Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within the Project Scope UID/CEC/00319/2013.

References

1. Fells, R.E.: Developing trust in negotiation. Employee Relations **15**(1), 33–45 (1993)
2. Castelfranchi, C., Falcone, R.: Trust is much more than subjective probability: Mental components and sources of trust. In: Proceedings of the 33rd Hawaii International Conference on System Sciences, HICSS 2000, Washington, DC, USA, vol. 6, p. 6008. IEEE Computer Society (2000)
3. Picard, R.W.: Computer learning of subjectivity. ACM Comput. Surv. **27**(4), 621–623 (1995)
4. Sherchan, W., Nepal, S., Paris, C.: A survey of trust in social networks. ACM Comput. Surv. 45(4), 47:1–47:33 (2013)
5. Carneiro, D., Novais, P., Andrade, F., Zeleznikow, J., Neves, J.: Using case-based reasoning and principled negotiation to provide decision support for dispute resolution. Knowledge and Information Systems **36**(3), 789–826 (2012)
6. Ross, W., LaCroix, J.: Multiple meanings of trust in negotiation theory and research: A literature review and integrative model. International Journal of Conflict Management **7**(4), 314–360 (1996)
7. Adali, S., Escriva, R., Goldberg, M.K., Hayvanovych, M., Magdon-Ismail, M., Szymanski, B.K., Wallace, W.A., Williams, G.: Measuring behavioral trust in social networks. In: 2010 IEEE International Conference on Intelligence and Security Informatics (ISI), pp. 150–152, May 2010
8. Gomes, M., Oliveira, T., Carneiro, D., Novais, P., Neves, J.: Studying the effects of stress on negotiation behavior. Cybernetics and Systems **45**(3), 279–291 (2014)
9. Pimenta, A., Carneiro, D., Novais, P., Neves, J.: Monitoring mental fatigue through the analysis of keyboard and mouse interaction patterns. In: Proceedings of Hybrid Artificial IntelligentSystems: 8th International Conference, HAIS 2013, Salamanca, Spain, September 11-13, pages 222–231. Springer, Heidelberg (2013)
10. Carneiro, D., Gomes, M., Novais, P., Neves, J.: Developing dynamic conflict resolution models based on the interpretation of personal conflict styles. In: Proceedings of Progress in Artificial Intelligence: 15th Portuguese Conference on Artificial Intelligence, EPIA 2011, Lisbon, Portugal, October 10-13, pp. 44–58. Springer, Heidelberg (2011)

A Persuasive Cognitive Assistant System

Angelo Costa, Stella Heras, Javier Palanca, Paulo Novais and Vicente Julián

Abstract In this paper, we present a persuasive recommendation module to be included into the iGenda framework. iGenda is a cognitive assistant that helps carereceivers and caregivers in the management of their agendas. The proposed new module will allow the system to select and recommend to the users the action that potentially best suits to his/her interests. The multi-agent approach followed by the iGenda framework facilitates an easy integration of these new features.

Keywords iGenda · AAL · Persuasive system · Persuasion · Recommendation

1 Introduction

Ambient Assisted Living (AAL) applications posses the capability to promote security and comfort for the users, trying to provide an integrated solution that connects several distinct devices to services to form a unique solution [1]. Current AAL applications can perceive the environment, monitoring events, and provide an adjusted and timely response that enables it to interact with the users. However, these features are not enough by themselves. The aim is to accommodate people, and people change their needs and behaviors. These important aspects are not taken into account in current AAL applications. Specifically, current AAL applications suffer from different problems to be directly used to persuade human behavior, mainly related with the social environment of each individual. Often the target audiences are large and

A. Costa(✉) · P. Novais
ALGORITMI Research Centre, University of Minho, Braga, Portugal
e-mail: {acosta,pjon}@di.uminho.pt

S. Heras · J. Palanca · V. Julián
Departamento de Sistemas Informáticos y Computación,
Universitat Politècnica de València, Valencia, Spain
e-mail: {sheras,jpalanca,vinglada}@dsic.upv.es

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_17

heterogeneous, and include users with wide-ranging goals, needs, and preferences. Thus, persuading the entire audience effectively with a one-size-fits-all persuasive intervention is difficult [2].

On the other hand, over the last few years, technology has been evolving in pursuit of persuading users and motivating them toward specific individual/collective behaviors. One of the most employed technology in order to try to persuade users is recommender systems. Current limitations of these systems are that traditional recommender systems base their recommendations on quantitative measures of similarity between the user's preferences and the current items to recommend (i.e. content-based recommenders [3]), between the user's profile and the profile of other users with similar preferences (i.e. collaborative filtering recommenders [4]) and on combinations of both (i.e. hybrid recommenders [5]). However, [6] has stated the inability of current recommender systems to use the large amount of qualitative and quantitative data available in AAL domains to empower recommendations. Usually, recommender systems do not provide an explanation about the reasoning process that has been followed to come up with specific recommendations. However, this does not follow current trends, where people trust recommendations more when the engine can provide reasons for them [7]. Thus, what is understood as a good recommendation is changing from the one that minimizes some error evaluation to the one that really makes people happier. Several studies demonstrated that the argumentation of the different recommendations has a strong impact to strengthen these recommendations, including direct feedback from the user and providing alternative options which better fits the user needs and expectations [6][8]. The personalized selection of arguments can streamline the persuasion task and make the desired behavior or attitude change easier to achieve.

In order to overcome these problems, our proposal is to embed a new persuasive layer into the iGenda platform [9, 10], taking into account aspects such as the generation of arguments that support tasks recommended by the AAL application to the users. Our work involves a contribution in the above commented areas, presenting a persuasive recommendation module to be used in iGenda in order to improve its system credibility.

2 The iGenda Framework

The iGenda is a cognitive assistant built upon a multi-agent system. It's purpose is to help care-receivers and caregivers by managing their agendas (create and reallocate events) and by presenting an easy-to-use visual interface for web and mobile usage, the platform is presented extensively in the papers [1, 9]. The aim of the iGenda is to be integrated on the users daily life and to become a constant helper, reminding the users of their appointments and events. It promotes active aging by introducing playful events on the users free time. Moreover, it serves as an visual interface to the users, showing new events and notifications and allowing the creation of new events.

It is designed to be used by the caregivers and the care-receivers alike, having each one a specially designed interface that accounts for their roles on the platform.

One of the issues with iGenda, and most of the AAL platforms, is that it does not gather the users input or opinion. Works by Holzinger el al. [11] and Lindley et al. [12] have showed that it is imperative that the users feel included and part of the decision process. It is hard to release the control of decision making after a life of making choices. In iGenda when an event is scheduled it is showed to the user but not explained why or give any option to decline or accept the event. The lack of reason and justification of why the events should be performed may lead to the withdrawal of the system by the users. A major advance would be to include persuasive methods that may provide a motive to that specific event scheduling and compel the users to attend to it.

In terms of the Free Time Manager (which promotes active aging through the scheduling of activities on the care-receivers free time) most of the activities scheduled by it are optional, but they greatly help the users to be active and promote an healthy life, but without a justification it is hard to put them in context or why they matter. Each time iGenda detects that a new action must be proposed to deal with the needs of the user, it selects a set of matching actions from its *free time events database*. In the previous version of iGenda, if the system was able to find more than one matching action to offer to the user, a single event was selected trough a biased random function; the selection leans towards the best ranked six events or an unbiased random event, the choice between the two functions is also random. However, the potential willingness of the user to accept a specific action (based on his/her current social context - i.e. the specific user, the specific caregiver, their relation, etc. - and the knowledge of similar past experiences) was not taken into account.

One of the main features is the modular architecture that the iGenda has. Its open connectivity allows the connection of new modules that follow the protocols [13, 14]. Being a multi-agent system, the iGenda allows easy integration of new features and the deployment and removal of agents at once [15].

3 Persuasion Module

In this section, we present the new persuasion module that we have added to the iGenda tool. The new persuasion module enhances the performance of iGenda, by allowing the system to select and recommend to the user the action that potentially best suits to his/her social context and interests. In addition, the user is presented with reasons that support the suitability of the action proposed. Thus, in this way the iGenda system tries to persuade the user to accept the action and to motivate his/her to put it into practice.

The persuasive module is based on the case-based argumentation framework for agent societies presented in [16]. Therefore, when iGenda has to select a specific event among a set, the system tries to create one argument (or more) to support each action. Then, an internal argumentation process takes part to decide the action that is better supported by its arguments.

3.1 Argumentation Framework

In this work, we have applied Agent-specific Argumentation Framework in an Agent Society ($AAFAS$) presented in [16] to determine which agent's argument attacks another agent's argument in an argumentation process performed in a society of agents and, in each case, which argument would defeat the other. To do that, we have to consider the values that arguments promote (the preferences of the users), the users' preference relations (preference orderings), and the dependency relations between agents (the relations that emerge from agent interactions or are predefined by the system).

Hence, our system models a society of agents with a set Ag of agents of the agent society S, a set of Rl agents' roles that have been defined in S, a set of D possible dependency relations in S, and a set of V values predefined in S. Thus, an agent specific argumentation framework for an agent society is a tuple $AAFAS = < Ag, Rl, D, V, A, Role, Dependency_S, val, Valpref_{ag_i} >$ where:

Definition 1 (Agent-specific AF for an Agent Society)

– Ag, Rl, D, and V are defined as the above elements of the agent society.
– A is the set of arguments of the argumentation framework.
– $Role(ag, a) : Ag \times A \rightarrow Rl$ is a function that assigns an agent the specific role that it plays (from its set of roles) when it has put forward a specific argument.
– $Dependency_S :<_D^S \subseteq Rl \times Rl$ defines a reflexive, transitive and asymmetric partial order relation over roles.
– $val(ag, a) : Ag \times A \rightarrow 2^V$ is a function that assigns an agent's argument the value(s) that it promotes.
– $Valpref_{ag_i} \subseteq V \times V$, defines an irreflexive, transitive and asymmetric relation $<_{ag_i}^S$ over the agent's ag_i values in the society S.

In our system, agents can play the role of *patients*, *caregivers* (which can be relatives, personal health assistants, friends, etc.), and *doctors*. We also consider the following dependency relations: (i) *Power*: when an agent has to accept a request from another agent because of some pre-defined domination relationship between them. For instance, in our agent society S, $Patient <_{Pow}^S Doctor$, and $Caregiver <_{Pow}^S Doctor$ since patients and caregivers must follow the guidelines recommended by their doctors; (ii) *Authorisation*: when an agent has committed itself to another agent for a certain service and a request from the latter leads to an obligation when the conditions are met. For instance, in S, $Patient <_{Auth}^S Caregiver$, if the patient has contracted the health assistant service that a caregiver offers; and (iii) *Charity*: when an agent is willing to accepts a request from another agent without being obliged to do so. For instance, in S, by default $Patient <_{Ch}^S Patient$, $Caregiver <_{Ch}^S Caregiver$ and $Doctor <_{Ch}^S Doctor$.

In addition, values represent preferences about the different types of activities that iGenda can recommend to the user. Concretely, in our system we have established the following typology of values and activities:

– *Motion* Values: which represent preferences for activities that are performed still or sitting (*motionless*), standing up or with little movement (*low-motion*), or those performed with physical effort (*motion*).
– *Location* Values: which represent preferences for activities that are performed indoors without movement (*homebound*), indoors with movement (*indoors*), outdoors with movement (*outdoors*), or outdoors with movement but that implies that the person has to be carried (*limited-outdoors*).
– *Social* Values: which represent preferences for activities that involve socialise with others (*social*), or not (*individual*).
– *Environmental Conditions* Values: which represent preferences for activities that only can be performed with good weather (*weather-dependent*), or not (*weather-independent*).
– *Health Conditions* Values: which represent preferences for activities that have immediate or direct impact on health (*mandatory*), or not (*not-mandatory*).

We have adapted one of the knowledge resources of this framework, the *argument-cases case base*, to use it as a persuasion resource for our system. This resource stores previous experiences and their final outcome in the form of case-based arguments, which can be retrieved and used later: 1) to generate new arguments that support each action; 2) to select the best action to recommend in view of past experiences; and 3) to store the new argumentation knowledge gained in each process, improving the system persuasion skills. The argument-cases are the main structure that we use to implement our framework and computationally represent arguments in agent societies. Table 1 shows an example of the structure of a specific argument-case in our system. The argument-cases have three main parts: the description of the *problem* that the case represents (i.e. the features that describe the situation where an action has to be recommended), the *solution* applied to this problem (i.e. the action recommended) and the *justification* why this particular solution was applied (i.e. specific features that match the situation, guidelines that have motivated the recommendation of a specific action, etc.). Therefore, an argument-case stores the information about a previous argument that an agent created to support the recommendation of a specific action.

Problem: The problem description stores the *premises* of the argument-case, which represent the context of the domain where the case was created. In addition, if we want to store an argument and use it to generate a persuasive argument in the future, the features that characterise the audience of the previous argument (the social context) must also be kept.

Therefore, we store in the argument-case the social information about the *proponent* of the argument, the *opponent* to which the argument is addressed, and the dependency relation established between the roles that these agents play. Thus, the proponent and opponent's features represent information about the agent that generated the argument and the agent that received it respectively. Concretely, for each agent the argument-case stores a unique *ID* that identifies it in the system and the *role* that the agent was playing when the argument was created. Moreover, if known,

Table 1 Structure of an Argument Case AC1

PROBLEM	Domain Context	Premises = {Sensors information, etc.}	
	Social Context	Proponent	ID = Doctor1 (D1)
			Role = Doctor
			$ValPref_{D1}$ = [indoors<outdoors]
		Opponent	ID = Caregiver1 (C1)
			Role = Patient
			$ValPref_{C1}$ = [outdoors<indoors]
		Dependency Relation = Power	
SOLUTION	Conclusion = $S1$ (Outdoors gymnastics activity)		
	Acceptability State = Unacceptable		
	Received Attacks	Distinguish Premises = $Rain$	
		Counter Examples = {AC2}	
JUSTIFICATION	Guidelines = {G.1.2}		

we also store the preferences of each agent over the pre-defined set of general values in the system. These preferences affect the persuasive power of the proponent's argument over the opponent's behaviour. Finally, the dependency relation between the proponent's and the opponent's roles is also stored.

Solution: In the solution part, the *conclusion* of the argument (the action that the argument-case supports) is stored. Moreover, the argument-case stores the information about the *acceptability state* of the argument at the end of the dialogue. This feature shows if the argument was deemed *acceptable*, *unacceptable* or *undecided* in view of the other arguments that were put forward. Regardless of the final acceptability state of the argument, the argument-case also stores the information about the possible *attacks* that the argument received. These attacks could represent the justification for an argument to be deemed unacceptable or else reinforce the persuasive power of an argument that, despite being attacked, was finally accepted. Argument-cases can store different types of attacks, depending on the type of argument that they represent: premises which value in the context where the argument was posed was different (or non-existent) than the value that it took when the argument-case was generated (*distinguish the case*) or argument-cases which premises also match the premises of the context where the argument was posed, but which conclusion is different than the conclusion of the case(s) used to generate the argument (*counter-examples*).

Justification: The justification part of the argument-case stores the information about the knowledge resources that were used to generate the argument represented by the argument-case. In our system, the justification stores the information of the health guidelines used to recommend a specific action for a specific user in a particular situation.

Following a CBR methodology, the knowledge resources of the agents' case-based argumentation system allow them to automatically generate, select and evaluate arguments. However, the complete argument management process (how agents generate, select and evaluate arguments by using the knowledge resources of their argumentation systems) is out of the scope of this paper. Also, the framework presented is flexible enough to represent different types of arguments and their associated information, but the value of some features on argument-cases could remain unspecified in specific domains. For instance, in some open MAS, the preferences over values of other agents could not be previously known. However, agents could try to infer the unknown features by using CBR adaptation techniques. Therefore, in our proposal, arguments that iGenda uses to are tuples of the form:

Definition 2 (Argument). $Arg = \{\phi, p, < SS >\}$, where ϕ is the conclusion of the argument, p is the preference value that the argument promotes and $< S >$ is a set of elements that justify the argument (the support set).

The support set $< SS >$ is the set of features (*premises*) that represent the context of the domain where the argument has been put forward (those premises that match the problem to solve and other extra premises that do not appear in the description of this problem but that have been also considered to draw the conclusion of the argument) and optionally, any knowledge resource used by the proponent to generate the argument (e.g. the health guidelines). On the other hand, the support set can also include any of the allowed attack elements of our framework. These are: *distinguishing premises*, or *counter-examples*.

Now, the concept of conflict between arguments defines in which way arguments can attack each other. There are two typical attacks studied in argumentation: *rebut* and *undercut*. In an abstract definition, rebuttals occur when two arguments have contradictory conclusions (i.e. if an argument a_1 supports a different conclusion for a problem description that includes the problem description of an argument a_2). Similarly, an argument undercuts other argument if its conclusion is inconsistent with one of the elements of the support set of the latter argument or its associated conclusion (i.e. if the conclusion drawn from the argument a_1 makes one of the elements of the support set of the argument a_2 or its conclusion non-applicable in the current recommendation situation). Thus, we can define the agent-specific defeat relation of $AAFAS$ as:

Definition 3 (Defeat). An agent's ag_1 argument $a_1 \in AAFAS$ that is put forward in the context of a society S $defeats_{ag_1}$ another agent's $ag_2 \in AAFAS$ argument a_2 iff $attack(a_1, a_2) \wedge (val(ag_1, a_1) <^S_{ag_1} val(ag_1, a_2) \notin Valpref_{ag_1}) \wedge (Role(ag_1) <^S_{Pow} Role(ag_2) \vee Role(ag_1) <^S_{Auth} Role(ag_2) \notin Dependency_S)$.

3.2 Example

To exemplify our framework, let us propose a simple scenario of an open MAS that represents a situation where iGenda has to schedule a health-care activity for a patient $P1$. Thus, iGenda retrieves from its activities database a potential activity $S1$ that was proposed by the patient's doctor $D1$ by following a health guideline G.1.2 and recommends an outdoors gymnastics activity that promotes the value *outdoors*. Also, it retrieves another activity $S2$ that was proposed by the patient's caregiver $C1$ by following another health guidelines and recommends an indoors training that promotes the value *indoors*. The doctor $D1$ has a value preference order $ValPref_{D1}$ = [indoors<outdoors] that promotes outdoors activities over indoors. On the contrary, let us assume that both the patient and his caregiver $C1$ have a value preference order $ValPref_{C1}$ = [outdoors<indoors] that represent their preference for indoors activities over outdoors. In addition, as established by our system, doctors have a power dependency relation over caregivers and patients that forces patients and caregivers to accept the recommendations provided by doctors. In this context, the iGenda system would internally generate a support argument (and its associated argument-case) for each possible activity to recommend as follows:

$A1 = \{S1, outdoors, < Patient1, SensorsInformation, G.1.2 >\}$ In view of the current context, Doctor $D1$ proposes the outdoors activity S1 following the health guidelines G.1.2. $A2 = \{S2, indoors, < Patient1, SensorsInformation, G.3.2 >\}$ In view of the current context, Caregiver $C1$ proposes the indoors activity S2 following the health guidelines G.3.2, which promotes the value indoors that matches the preference of the patient to do indoors activities over outdoors.

Up to this point, iGenda would have to schedule the activity $S1$ for the patient, since the *power* dependency relation of doctors over caregivers and patients will prevail, $A1$ *rebuts* $A2$, and $A2$ would be defeated by $A1$. However, let us assume that the day is rainy. With this new information, the system would be able to find and retrieve the argument-case $AC2$, which stores the information about a similar situation where an indoors activity $S2$ was proposed by the doctor due to the bad weather conditions. Then, the system will generate the following counter-example for the argument $A1$:

$A2.1 = \{S2, indoors, < Patient1, Rain, AC2 >\}$ In view of the current context, Caregiver $C1$ proposes the indoors activity S2 following the health guidelines G.3.2, which promotes the value indoors, matches the preference of the patient to do indoors activities over outdoors, and takes into account the current rainy weather conditions.

Clearly, $A2.1$ *undercuts* $A1$, since $A1$ did not take into account the new premise *rain* and hence, $A1$ does not longer match the current situation of the patient. Therefore, $A1$ would be defeated and iGenda would recommend the activity *S2* for the patient. If no more activities and their associated support arguments are posed, iGenda will finish the argumentation process and will store the new argument-case $AC1$ (as represented in 1). This new argument-case improves the recommendation skills of the system by representing the knowledge gained in the above argumentation process. Then, from now on and unless other argumentation process updates the information of the argument-cases case-base, the outdoor activity $S1$ would never be proposed again in rainy weather conditions (although it matches the preferences of the doctor). In addition, the argument-cases can be used to prompt persuasion messages for the user of iGenda. For instance, the system may create persuasion messages as "iGenda recommends you to perform activity S1 since you prefer indoor activities", "iGenda recommends you to perform activity S1 since is rainy" or "iGenda recommends you to perform activity S1 since it follows the guidelines of your doctor when rains".

4 Conclusions

This paper has presented a new persuasive oriented module to be included into the iGenda framework. The persuasive module is a case-based argumentation approach which allows iGenda not only to select a specific event among a set, but also it allows the system to create one argument (or more) to support each possible event. To do this, an argumentation process takes part into the system to decide the action that is better supported by its arguments. This process enhances the response given by iGenda to the user because the selected action is presented with reasons that supports it. Thus, the proposed solution tries to persuade the user to accept the action and to put it into practice.

Acknowledgements This work has been partially supported by the projects TIN2015-65515-C4-1-R and TIN2014-55206-R of the Spanish government, by the grant program for the recruitment of doctors for the Spanish system of science and technology (PAID-10-14) of the Universitat Politècnica de València, and by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the projects UID/CEC/00319/2013 and Post-Doc scholarship SFRH/BPD/102696/2014 (A. Costa).

References

1. Costa, Â., Castillo, J.C., Novais, P., Fernández-Caballero, A., Simoes, R.: Sensor-driven agenda for intelligent home care of the elderly. Expert Systems with Applications **39**(15), 12192–12204 (2012)
2. Berkovsky, S., Kuflik, T., Ricci, F.: Mediation of user models for enhanced personalization in recommender systems. User Modeling and User-Adapted Interaction **18**(3), 245–286 (2007)
3. Pazzani, M.J., Billsus, D.: Content-based recommendation systems. In: The Adaptive Web, pp. 325–341. Springer (2007)
4. Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative filtering recommender systems. In: The Adaptive Web, pp. 291–324. Springer (2007)
5. Burke, R.: Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction **12**(4), 331–370 (2002)
6. Chesñevar, C., Maguitman, A.G., González, M.P.: Empowering recommendation technologies through argumentation. In: Argumentation in Artificial Intelligence, pp. 403–422. Springer (2009)
7. Linden, G., Hong, J., Stonebraker, M., Guzdial, M.: Recommendation algorithms, online privacy, and more. Communications of the ACM **52**(5), 10 (2009)
8. Heras, S., Navarro, M., Botti, V., Julián, V.: Applying dialogue games to manage recommendation in social networks. In: Lecture Notes in Computer Science, pp. 256–272. Springer (2010)
9. Costa, A., Novais, P., Corchado, J.M., Neves, J.: Increased performance and better patient attendance in an hospital with the use of smart agendas. Logic Journal of IGPL **20**(4), 689–698 (2011)
10. Costa, A., Novais, P., Simoes, R.: A caregiver support platform within the scope of an ambient assisted living ecosystem. Sensors **14**(3), 5654–5676 (2014)

11. Holzinger, A., Ziefle, M., Röcker, C.: Human-computer interaction and usability engineering for elderly (HCI4Aging): introduction to the special thematic session. In: Lecture Notes in Computer Science, pp. 556–559. Springer (2010)
12. Lindley, S., Wallace, J.: Placing in age. ACM Transactions on Computer-Human Interaction **22**(4), 1–39 (2015)
13. Andrade, F., Novais, P., Machado, J., Neves, J.: Contracting agents: legal personality and representation. Artificial Intelligence and Law **15**(4), 357–373 (2007)
14. Andrade, F., Neves, J., Novais, P., Machado, J., Abelha, A.: Legal security and credibility in agent based virtual enterprises. In: IFIP — The International Federation for Information Processing, pp. 503–512. Springer (2005)
15. Castillo, J.C., Carneiro, D., Serrano-Cuerda, J., Novais, P., Fernández-Caballero, A., Neves, J.: A multi-modal approach for activity classification and fall detection. International Journal of Systems Science **45**(4), 810–824 (2013)
16. Heras, S., Botti, V., Julián, V.: Argument-based agreements in agent societies. Neurocomputing **75**(1), 156–162 (2012)

Preliminary Study of Classifier Fusion Based Indoor Positioning Method

Yuki Miyashita, Mahiro Oura, Juan F. De Paz, Kenji Matsui, Gabriel Villarrubia and Juan M. Corchado

Abstract Indoor positioning technology is commercially available now, however, the positioning accuracy is not sufficient in the current technologies. Currently available indoor positioning technologies differ in terms of accuracy, costs and effort, but have improved quickly in the last couple of years. It has been actively conducted research for estimating indoor location using RSSI (Received Signal Strength Indicator) level of Wi-Fi access points or BLE (Bluetooth Low Energy) tags. WiFi signal is commonly used for the indoor positioning technology. However, It requires an external power source, more setup costs and expensive. BLE is inexpensive, small, have a long battery life and do not require an external energy source. Therefore, by adding some BLE tags we might be able to enhance the accuracy inexpensive way. In this paper, we propose a new type of indoor positioning method based on WiFi-BLE fusion with Fingerprinting method. WiFi RSSI and BLE RSSI are separately processed each one by a Naive Bayes Classifier. Then, Multilayer Perceptron(MLP) is used as the fusion classifier. Preliminary experimental result shows 2.55m error in case of the MLP output. Since the result is not as good as the ones using conventional method, further test and investigation needs to be performed.

Keywords Indoor positioning · Classifier fusion · Wi-Fi · BLE · Fingerprint

Y. Miyashita(✉) · K. Matsui
Department of Engineering, Osaka Institute of Technology, 5-16-1 Omiya, Asahi Ward, Osaka 535-0002, Japan
e-mail: m1m15341@st.oit.ac.jp, kenji.matsui@oit.ac.jp

M. Oura
Department of Information Science and Technology, Osaka Institute of Technology, 1-79-1 Kitayama, Hirakata City, Osaka 573-0196, Japan
e-mail: m1m15a05@st.oit.ac.jp

J.F. De Paz · G. Villarrubia · J.M. Corchado
BISITE Research Group, University of Salamanca, Edificio I+D+I, 37008 Salamanca, Spain
e-mail: {fcofds,gvg,corchado}@usal.es

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_18

1 Introduction

In case of outdoor positioning, satellite-based GPS positioning works very well. However, indoor positioning [2] [5] is not so straight forward, and that area is glowing research fields in mobile computing because of the popularization of mobile devices, like smartphones, tablets. The technologies currently most used in the development of Real-Time Location System are, RFID (Radio Frequency IDentification) [10], Wi-Fi [6] or BLE (Bluetooth Low Energy) [4] [7]. Indoor positioning technology is commercially available now, however, the positioning accuracy is not sufficient in current technologies. Currently available indoor positioning technologies differ in terms of accuracy, costs and effort, but have improved quickly in the last couple of years. WiFi signal is commonly used for the indoor positioning technology. However, WiFi equipment requires an external power source and more expensive setup costs. WiFi signal is strong and it can cover relatively wide area. Bluetooth Low Energy (BLE) is one of the latest technologies. It is called BLE beacons (or iBeacons) that are inexpensive, small, have a long battery life and do not require an external energy source. Both WiFi and BLE technologies bring a couple of advantages: easy deployment and are integrated in most of the current electronic devices. Thus, it has been actively conducted research for estimating indoor location using RSSI (Received Signal Strength Indicator) level of Wi-Fi access points or BLE tags.

This paper reports a preliminary study of WiFi-BLE mixed indoor positioning engine, which can provide location information of people or objects inside a building with possibly better accuracy and better cost performance. The use of these technologies for location system is mainly based on intensity maps constructed from RSSI levels in different zones. The maps are used as a basis of obtaining the locations. The classifiers use the data of the maps and the data obtained from the devices to determine the position of a person inside a building.

In this article, Section 2 describes our proposed method, Section 3 presents the detailed points of the experiment. Section 4 summarizes the experimental results, and in Section 5, we discuss about our next steps, finally, we conclude our report.

2 Proposed Method

There are various methods in terms of indoor positioning. In the study written by Gabriel [1], he used WiFi RSSI data with Bayesian network based classifier to get the estimated position data. The intensity map has the data structure shown in Figure1. Each data map contains the information of all the WiFi access point MAC addresses, RSSI values, as well as the coordinates (x, y). The RSSI values are considered to be the distance from the WiFi access points. The more data we use, the more processing time we need.

Our objectives are: 1) increase the positioning accuracy, 2) increase the processing speed, and 3) increase the cost-performance. In order to meet those objectives, first, we added more beacons to make the position estimation accurate. For those beacons, we chose simple BLE tags to reduce the additional cost. Then, we introduced the combination of Bayes classifier and Multi-layer Perceptron (MLP) in order to increase the processing speed without losing the accuracy. Similar study was reported by Javier [3]. He proposed the method using fusion classifier.

Figure 2 shows the basic building block of the proposed fusion classifier, where $\sum f(pos_i)$ is the expected location coordinates. The advantages of using BLE tags are not only improving the cost performance, but we might be able to implement many applications associated with BLE tags.

X	*Y*	*MACAddress*	*RSSI*	*MACAddress*	*RSSI*	...	...

Fig. 1 Data format of the intensity map

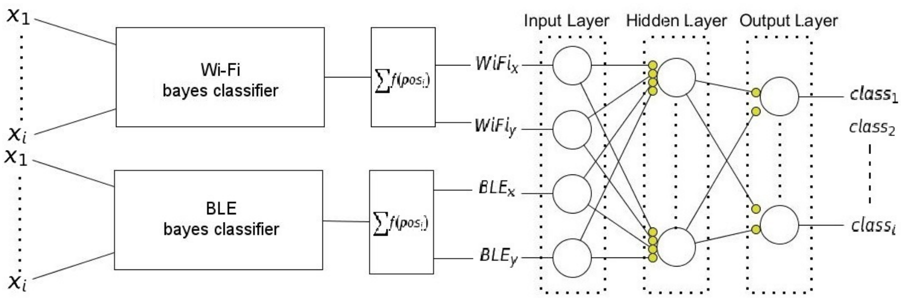


Fig. 2 Block diagram of the proposed fusion classifier

3 Experiment

As for the test environment, we used our office shown in Figure 3. Three square symbols in that figure are WiFi access points, round ones are BLEs, and 100 fingerprints are indicated by X marks. That office is 15 × 20m, approximately $300m^2$. There are around 30 additional WiFi access points in the building, and we allocated 8 BLE tags in the office. For the training data collection, one smart phone was utilized to measure the RSSI and other data. We measured the training data from 100 locations in the room as total. At each location, RSSI values were measured 10 times. That means 100 fingerprints were collected. Those collected data were used to train the Bayes classifiers. The output from the Bayes classifier is the estimated location coordinates (x_i,y_i). Both those estimated data and actual measured data were used to tune the MLP.

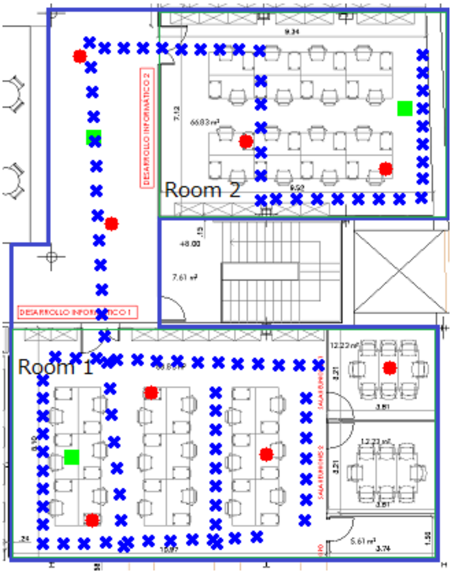


Fig. 3 Environment of evaluation

In this study, our proposed indoor positioning system is based on the fingerprinting [7] on a map of intensities. In terms of the accuracy evaluation, we compared four different positioning methods as shown below. The overall diagram is shown in Figure 4.

1. Bayes classifier with WiFi - RSSI
2. Bayes classifier with BLE - RSSI
3. Bayes classifier with both WiFi - RSSI and BLE - RSSI
4. MLP with Method 1 output and Method 2 output.

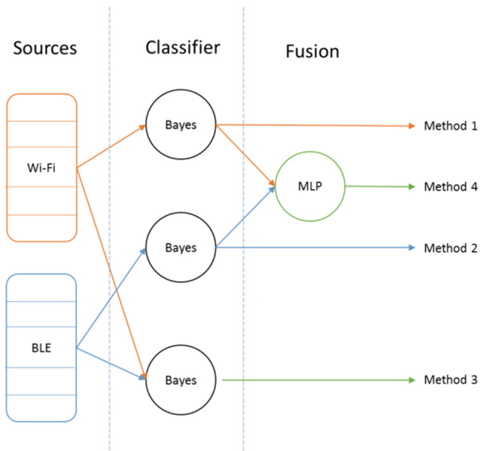


Fig. 4 Block diagram of the experiment system

For the experiment, we used around 30 WiFi signals in the building as well as the three WiFi stations in the room. Those additional access points are the originally installed ones in the building. As for the testing data, we obtained 50 more measurement data, and used for the experiment.

4 Experimental Results

The experimental results are shown in Table 1. The average error shows that the proposed method (method 4) was not as good as other methods. In terms of the BLE tags, both method 2 and method 3 are slightly better than the WiFi based location estimation. More detailed evaluation with various conditions needs to be performed.

Table 1 Errors of estimating [m]

Method	*Max Error*	*Minimum Error*	*Average Error*
1	6.159	0.173	2.525
2	6.367	0.198	2.254
3	6.527	0.240	2.363
4	5.898	0.292	2.555

5 Discussion

There are many issues we need to work further. For instance, how to stabilize the location estimation is very important challenge. Since we did not address about the WiFi intensity signal fluctuation associated with the equipment characteristics, variation in power source, and various room environment, some of the promising directions of improvement are: (1) Updating the fingerprint data based on carefully observed signal intensity behavior. The fingerprints are normally influenced by environmental factors, the brands and models of the devices. (2) Eliminating negatively affect devices by observing intensity level and fluctuation. (3) Applying the each probability distribution functions into the Bayes classifiers.

6 Conclusion

In this paper, we provided a preliminary study of indoor positioning estimation using WiFi-BLE fusion method. As for the classifiers, we tested the combination of Bayes classifer and MLP. We evaluated the overall fingerprint matching performance between the proposed method and conventional method. The average error shows that the proposed method was not as good as other methods. In terms of the BLE tags, both method 2 and method 3 are slightly better than the WiFi based location estimation. Since both WiFi and BLE signal intensity is normally influenced by environmental factors and device characteristics, more detailed evaluation with various conditions needs to be performed.

References

1. Villarubia, G., Rubio, F., De Paz, J.F., Bajo, J., Zato, C.: Applying classifiers in indoor location system. In: Trends in Practical Applications of Agents and Multiagent Systems, pp. 53–58 (2013)
2. Kaemarungsi, K., Kaemarungsi, K.: Modeling of indoor positioning systems based on location fingerprinting. In: INFOCOM 2004. Twenty-Third Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 2, pp. 1012–1022 (2004)
3. Bajo, J., De Paz, J.F., Villarrubia, G., Corchado, J.M.: Self-Organizing Architecture for Information Fusion in Distributed Sensor Networks. International Journal of Distributed Sensor Networks **2015** (2015). Hindawi Publishing Corporation
4. Major League Baseball: The Official Ballpark App. http://mlb.mlb.com/mobile/ballpark/ (accessed January 20, 2016)
5. Villarrubia, G., De Paz, J.F., Bajo, J., Corchado, J.M.: Monitoring and Detection Platform to Prevent Anomalous Situations in Home Care. Sensors **14**, 9900–9921 (2014)
6. Zaruba, G.V., Huber, M., Kamangar, F.A., Chlamtac, I.: Indoor location tracking using RSSI readings from a single Wi-Fi access point. Wireless Networks **13**, 221–235 (2007)
7. Hui, L., Darabi, H., Banerjee, P., Jing, L.: Survey of Wireless Indoor Positioning Techniques and Systems. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews **37**, 1067–1080 (2007)
8. Madigan, D., Einahrawy, E., Martin, R.P., Ju, W.H., Krishnan, P., Krishnakumar, A.S.: Bayesian indoor positioning systems. In: INFOCOM 2005, 24th Annual Joint Conference of the IEEE Computer and Communications Societies Proceedings IEEE, vol. 2, pp. 1217–1227 (2005)
9. Riedmiller, M.: Advanced supervised learning in multi-layer perceptrons – From backpropagation to adaptive learning algorithms. Computer Standards & Interfaces **16**, 265–278 (1994)
10. Sala, M., Guzman, R., Egea, W.: Using neural networks and Active RFID for indoor location services. In: 2010 European Workshop on Smart Objects: Systems, Technologies and Applications (RFID Sys Tech), pp. 1–9 (2010)

Providing Wellness Services Using Real Time Analytics

Daniel Araújo, André Pimenta, Davide Carneiro and Paulo Novais

Abstract Data has increased in a large scale in various fields leading to the coin of the term Big Data. Big data is mainly used to describe enormous datasets that typically includes masses of unstructured data that may need real-time analysis. As human behaviour and personality can be captured through human-computer interaction a massive opportunity opens for providing wellness services. Through the use of interaction data, behavioral biometrics can be obtained. The usage of biometrics has increased due to several factors such as the rise of power and availability of computational power. One of the challenges in this kind of approaches has to do with handling the acquired data. The growing volumes, variety and velocity brings challenges in the tasks of pre-processing, storage and providing analytics. In this sense, the problem can be framed as a Big Data problem. In this work it is intended to provide an architecture that accommodates the data pipeline of data generated by human-computer interaction, providing real time data analytics on behavioral biometrics.

Keywords Real-time analytics · Big data · NoSQL databases · Behavioral biometrics

1 Introduction

A large amount of data is created every day by the interactions of billions of people with computers, wearable devices, GPS devices, smart phones, and medical devices. In a broad range of application areas, data is being collected at unprecedented scale [1].

D. Araújo(✉) · A. Pimenta · D. Carneiro · P. Novais
Algoritmi Center/Department of Informatics, Minho University, Braga, Portugal
e-mail: daniel.ara.6@gmail.com, {apimenta,dcarneiro,pjon}@di.uminho.pt

D. Carneiro
CIICESI, ESTGF, Polytechnic Institute of Porto, Felgueiras, Portugal

© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_19

Not only the volume of data is growing, but also the variety (range of data types and sources) and velocity (speed of data in and out) of data being collected and stored. These are known as the 3V's of Big data, enumerated in a research report published by Gartner [2].

Big Data refers to things one can do at a large scale that cannot be done at a smaller one: to extract new insights or create new forms of value, in ways that change markets, organizations, the relationship between citizens and governments, and more [3]. Despite still being somewhat an abstract concept it can be clearly said that Big Data encomprises the a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling the high-velocity capture, discovery and analysis [4].

1.1 Human-Computer Interaction

With the advent of the Internet of Things the number of devices that are connected is increasing every. Consequently, the number of interactions that can generate data is growing as well. Humans tend to show their personality or their state through their actions, even in an unconscious way. Facial expressions and body language, for example, have been known as a gateway for feelings that result in intentions. The resultant actions can be traced to a certain behavior. Therefore, it is safe to assume that a human behavior can be outlined even if the person does not want to explicitly share that information.

The interaction with computers and other technological devices can provide dataset containing records that are relative to unconscious behaviors. The rhythm at which a person types on a keyboard or the movement of the mouse changes when the individual becomes fatigued or under severe stress, as it was established in [5, 6]. Moreover, the interaction with smartphones can also provide analogous patterns from diverse sources including a touchscreen (that provides information about touches, their intensities, their area or their duration), gyroscopes, accelerometers, among others.

2 NoSQL for Real Time Analytics

2.1 Real Time Analytics

In the spectrum of analytics two extremes can be identified. On one end of the spectrum there is batch analytical applications, which are used for complex, long-running analyses. Generally, these have slower response times (hours or days) and lower requirements for availability. Hadoop-based workloads are an example of batch analytical applications. On the other end of the spectrum sit real-time analytical applications. Real-time can be considered from the point of view of the data or from the point of view of the end-user. The earlier translates into the ability of processing data as it arrives, making it possible to aggregate data and extract trends about the actual situation of the system (streaming analytics). The former refers to the ability to process data with low latency (processing huge amount of data with the results being available for the end user almost in real-time) making it possible, for example, to to provide recommendations for an user on a website based on its history or to do unpredictable, ad hoc queries against large data sets (online analytics).

Regarding stream processing the main problems are related to: Sampling Filtering, Correlation, Estimating Cardinality, Estimating Quantiles, Estimating Moments, Finding Frequent Elements, Counting Inversions, Finding Subsequences, Path Analysis, Anomaly Detection Temporal Pattern Analysis, Data Prediction, Clustering, Graph analysis, Basic Counting and Significant Counting. The main applications are A/B testing, set membership, fraud detection, network analysis, traffic analysis, web graph analysis, sensor networks and medical imaging ([7]).

According to [7] these are the most well-known streaming open source tools:

S4 Real-time analytics with a key-value based programming model and support for scheduling/message passing and fault tolerance.

Storm The most popular and widely adopted real-time analytics platform developed at Twitter.

Millwheel Google's proprietary real-time analytics framework that provides exact once semantics.

Samza Framework for topology-less real-time analytics that emphasizes sharing between groups.

Akka Toolkit for writing distributed, concurrent and fault tolerant applications.

Spark Does both offline and online analysis using the same code and same system.

Flink Fuses offline and online analysis using traditional RDBMS techniques.

Pulsar Does real-time analytics using SQL.

Heron Storm re-imagined with emphasis on higher scalability and better debuggability.

Online analytics, on the other hand, are designed to provide lighter-weight analytics very quickly. The requirements of this kind of analytics are low latency and high availability. In the Big Data era, OLAP (online analytical processing [8]) and traditional ETL processes are too expensive. Particularly, the heterogeneity of the data sources makes it difficult the definition of rigid schemas, making model-driven insight difficult hard. In this paradigm analytics are needed in near real time in order to support operational applications and their users. This includes applications from social networking news feeds to analytics, from real-time ad servers to complex CRM applications.

2.2 NoSQL - Not only SQL

Conventional relational databases have proven to be highly efficient, reliable and consistent in terms of storing and processing structured data [9]. However, regarding the 3 V's of big data the relational model has several shortcomings. Companies like Amazon, Facebook and Google started to work on their own data engines in order to deal with their Big Data pipeline, and this trend inspired other vendors and open source communities to do similarly for other use cases. As Stonebraker argues in [10] the main reasons to adopt NoSQL databases are performance (the ability to manage distributed data) and flexibility (to deal with semi-structured or unstructured data that may arise on the web) issues.

A mapping between Big Data characteristics (the 3V's) and NoSQL features can be established. NoSQL data stores can manage large volumes of data by enabling data partitioning across many storage nodes and virtual structures, overcoming traditional infrastructure constraints (and ensuring basic availability). By compromising on ACID (Atomicity, Consistency, Isolation, Durability ensured by RDBMS in database transactions) properties NoSQL opens the way for less blocking between user queries. The alternative is the BASE system [11] that translates to basic availability, soft state and eventual consistency. By being basically available the system is guaranteed to be mostly available, in terms of the CAP theorem. Eventual consistency indicates that given that the system does not receive input during an interval of time, it will become consistent. The soft state propriety means that the system may change over time even without input.

According to [12], the key characteristics that generally are part of NoSQL systems are, the ability to horizontally scale CRUD operations throughput over many servers, the ability to replicate and to distribute (i.e.,partition or shard) data over many servers, a simple call level interface or protocol (in contrast to a SQL binding), a weaker concurrency model than the ACID transactions of most relational (SQL) database systems, efficient use of distributed indexes and RAM for data storage, and the ability to dynamically add new attributes to data records.

However, the systems differ in many points, as the funcionality ranges from a simple distributed hashing (as supported by memcached[1], an open source cache), to highly scalable partitioned tables (as supported by Google's BigTable [13]). NoSQL data stores come in many flavors, namely data models, and that permits to accommodate the data variety that is present in real problems.

Key-Value Stores. A Key-value DBMS can only perform two operations: store pairs of keys and values, and retrieve the stored values given a key. These kind of systems are suitable for applications with simple data models that require a resource-efficient data store like, for example, embedded systems or applications that require a high performance in-process database. Redis and Memcahced are popular examples of this kind of database.

1 http://memcached.org

Document-Oriented Databases. These kind of data stores are designed to store and manage documents. Typically, these documents are encoded in standard data exchange (such as XML, JSON, YAML, or BSON). These kind of stores allow nested documents or lists as values as well as scalar values, and the attribute names are dynamically defined for each document at runtime. A single column can hold hundreds of attributes (in an analogy to the relational model), and the number and type of attributes recorded can vary from row to row, since its schema free. Unlike key-value stores, these kind of stores allow the search on both keys and values, support complex keys and secondary indexes. MongoDB and CouchDB are DBMS that function in this paradigm.

Column-Oriented Databases. Column-oriented databases are the kind of data store that most resembles the relational model on a conceptual level. They retain notions of tables, rows and columns, creating the notion of a schema, explicit from the client's perspective. In this approach, rows are split across nodes through sharding on the primary key. They typically split by range rather than a hash function. This means that queries on ranges of values do not have to go to every node. Columns of a table are distributed over multiple nodes by using "column groups". Rows are grouped into collections (tables), and an individual row's attributes can be of any type. For applications that scan a few columns of many rows, they are more efficient, because this kind of operations lead to less loaded data than reading the whole row. Apache Cassandra and Apache HBase are examples of this kind of data store.

Graph-Oriented Databases. Graph databases are data stores that employ graph theory concepts. In this model, nodes are entities in the data domain and edges are the relationship between two entities. Nodes can have properties or attributes to describe them. These kind of systems are used for implementing graph data modeling requirements without the extra layer of abstraction for graph nodes and edges. This means less overhead for graph-related processing and more flexibility and performance. Neo4J is the most popular graph-oriented database.

Comparative Evaluation of NoSQL Databases. As it was presented, there are several options when it comes the time to choose a NoSQL database, and the different categories and architectures serve different purposes. Although four categories were presented, only two of them are adequate for the purposes of this work. Regarding support for complex queries column-oriented and document-oriented data store systems are more adequate than key-value stores (e.g. simple hash tables) and graph databases (which are ideal for situations that are modeled as graph problems). Considering the last presented fact a comparison is presented at [14], where several DBMS are classified in a 5-point scale (Great, good, average, mediocre and bad) regarding a set of quality attributes.

3 Wellness Services Using Real Time Analytics

Gathering metrics on people's behaviours and providing tools for visualization, particularly real time analytics, enables decision making and data-driven actions concerning well being of individuals. The trend for data collection regarding sensing on humans is growing and the perspective is for this trend to keep strong, giving the expected growth of IoT (Internet of Things). Moreover, Big Data tools and techniques enable for this to be done at a large scale without compromising performance and availability.

According to [5], by recording the data from the keyboard and mouse movements it is possible to metrics that enable the prediction of fatigue levels. The captured records contain fifteen values (represented as doubles) that are a result of applying data redundancy techniques (i.e. aggregation of collected data by calculating values such as mean and variance on the very frequently collected values) and additionally contain a timestamp. Therefore, each record needs (15 times 8 bytes, the MongoDB double size plus 8 bytes relative to the timestamp, and 8 bytes relative to two keys that refer the task and user) 136 bytes of storage space. These records are produced every five minutes for each user of the system. As a user is expected to be around eight hours per day (13056 bytes, 12.75 Kbytes) interacting with its desktop/laptop a prediction about the data volumes that need real time processing can be made (see figure 1).

Table 1 Data growth projections.

	1 user	**100 users**	**10000 users**	**1000000 users**
5 minutes	136 bytes	13.28 Kbs	1.297 Mbs	129.7 Mbs
1 day	12.75 Kbs	1.245 Mbs	124.5 Mbs	12.159 Gbs
1 week	89.25 Kbs	8.716 Mbs	871.6 Mbs	85.115
1 month	382.5 Kbs	37.354 Mbs	3.648 Gbs	364.8 Gbs
1 year	4.545 Mbs	454.5 Mbs	44.382 Gbs	4.438 Tbs

As it is shown in figure 1 the architecture of the desired system is divided in three major components. The raw data is generated in the devices, then pre-processed (by redundancy elimination) and stored locally whenever possible (as in smartphones and personal computers) in a SQLite database. Then data is synchronized with the web servers in the cloud. The target database is MongoDB (object-oriented DB). MongoDB[2] is a database that is half way between relational and non-relational systems. It provides indexes on collections, it is lockless and provides a query mechanism. MongoDB provides atomic operations on fields like relational systems MongoDB supports automatic sharding by distributing the load across many nodes with automatic failover and load balancing, on the other hand CouchDB achieves scalability through asynchronous replication. MongoDB supports replication with automatic failover and recovery. The data is stored in a binary JSON-like format called BSON

---

[2] https://www.mongodb.com

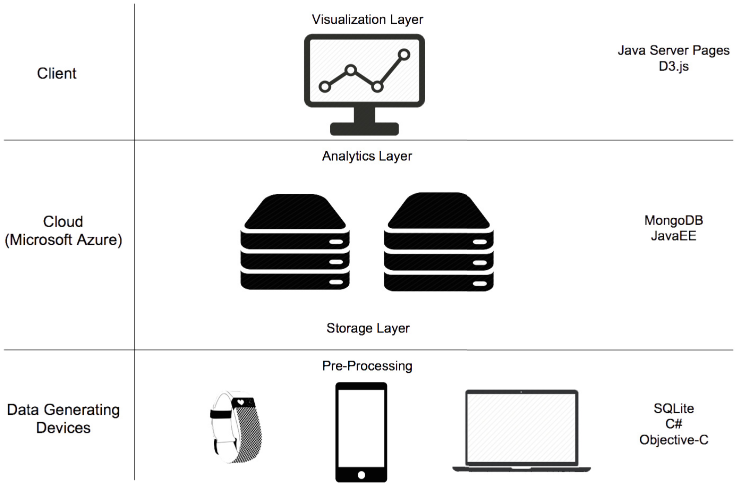


Fig. 1 Architecture of the real time analytics system.

that supports boolean, integer, float, date, string and binary types. The communication is made over a socket connection (in CouchDB it is made over an HTTP REST interface).

MongoDB is actually more than a data storage engine, as it also provides native data processing tools: MapReduce[3] and the Aggregation pipeline[4]. Both the aggregation pipeline and mapreduce can operate on a sharded collection (partitioned over many machines, horizontal scaling). These are powerful tools for performing analytics and statistical analysis in real-time, which is useful for ad-hoc querying, pre-aggregated reports, and more. MongoDB provides a rich set of aggregation operations that process data records and return computed results, using this operations in the data layer simplifies application code and limits resource requirements. The visualization layer (as an web app) is developed on Java technology and uses the D3 library for graphics and diagrams. Regarding fault tolerance MongoDB provides master-slave replication and replica sets[5]. Nowadays, replica sets are recommended for most use cases. The standard (and minimum) number of replicas in a set is three: one being the primary (the only one with writes allowed), and two secondaries (can become the primary in an election), since an odd number of members ensures that the replica set is always able to elect a primary. Another aspect in the data architecture that must be addressed is related to the storage engines. As the there is no longer a universal database storage technology capable of powering every type of application built by the business, MongoDB provides pluggable storage engines, namely

3 https://docs.mongodb.org/manual/core/map-reduce/

4 https://docs.mongodb.org/manual/core/aggregation-pipeline/

5 https://docs.mongodb.org/manual/replication/

WiredTiger and MMAPv1. Multiple storage engines can co-exist within a single MongoDB replica set, making it easy to evaluate and migrate engines. Running multiple storage engines within a replica set can also simplify the process of managing the data lifecycle. WiredTiger (default storage engine starting in MongoDB 3.2) will provide significant benefits in the areas of lower storage costs, greater hardware utilization, and more predictable performance[6] and, consequently should be used in this system.

4 Conclusions

In this work we presented an approach for handling the Big Data problems that may arise from a large scale wellness application. The rate of data arriving, the volume of the data and the need for real time aggregations (for data visualization) represent the data requirements. The main challenges are related to the databases architecture, particularly the storage and analytical layers. The database management system and the replica set architecture have been presented. The analytical engine is the Mongo Aggregation Framework, a feature of MongoDB that provides a pipeline for low level analytics operations. The deployment of this data architecture will allow us to test the ability of this kind of services to scale.

The door for large scale wellness services is opened, and big data techniques appear as one of the most useful resources for this kind of services to succeed. By aligning this trend with the advances on machine learning distributed systems and Internet of things applications, a future for diverse wellness services can be sighted.

Acknowledgement This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013.

References

1. Bertino, E., Bernstein, P., Agrawal, D., Davidson, S., Dayal, U., Franklin, M., Gehrke, J., Haas, L., Halevy, A., Han, J., et al.: Challenges and opportunities with big data (2011)
2. Gartner: What is big data? http://www.gartner.com/it-glossary/big-data (accessed: 2015-12-20)
3. Mayer-Schönberger, V., Cukier, K.: Big Data: A Revolution that will Transform How We Live, Work, and Think. Houghton Mifflin Harcourt (2013)
4. Gantz, J., Reinsel, D.: Extracting value from chaos. IDC iview **1142**, 9–10 (2011)
5. Pimenta, A., Carneiro, D., Novais, P., Neves, J.: Monitoring mental fatigue through the analysis of keyboard and mouse interaction patterns. In: Hybrid Artificial Intelligent Systems, pp. 222–231. Springer (2013)
6. Carneiro, D., Castillo, J.C., Novais, P., Fernández-Caballero, A., Neves, J.: Multimodal behavioral analysis for non-invasive stress detection. Expert Systems with Applications **39**(18), 13376–13389 (2012)

[6] https://docs.mongodb.org/manual/core/storage-engines/

7. Kejariwal, A., Kulkarni, S., Ramasamy, K.: Real time analytics: algorithms and systems. Proceedings of the VLDB Endowment **8**(12), 2040–2041 (2015)
8. Chaudhuri, S., Dayal, U.: An overview of data warehousing and olap technology. ACM Sigmod record **26**(1), 65–74 (1997)
9. Khazaei, H., Fokaefs, M., Zareian, S., Beigi-Mohammadi, N., Ramprasad, B., Shtern, M., Gaikwad, P., Litoiu, M.: How do i choose the right nosql solution? a comprehensive theoretical and experimental survey. Submitted to Journal of Big Data and Information Analytics (BDIA) (2015)
10. Stonebraker, M.: Sql databases v. nosql databases. Communications of the ACM **53**(4), 10–11 (2010)
11. Pritchett, D.: Base: An acid alternative. Queue **6**(3), 48–55 (2008)
12. Cattell, R.: Scalable sql and nosql data stores. ACM SIGMOD Record **39**(4), 12–27 (2011)
13. Chang, F., Dean, J., Ghemawat, S., Hsieh, W.C., Wallach, D.A., Burrows, M., Chandra, T., Fikes, A., Gruber, R.E.: Bigtable: A distributed storage system for structured data. ACM Transactions on Computer Systems (TOCS) **26**(2) (2008)
14. Lourenço, J.R., Cabral, B., Carreiro, P., Vieira, M., Bernardino, J.: Choosing the right nosql database for the job: a quality attribute evaluation. Journal of Big Data **2**(1), 1–26 (2015)

Smart Cities Simulation Environment for Intelligent Algorithms Evaluation

Pablo Chamoso, Juan F. De Paz, Sara Rodríguez and Javier Bajo

Abstract This paper presents an adaptive platform that can simulate the centralized control of different smart city areas. For example, public lighting and intelligent management, public zones of buildings, energy distribution, etc. It can operate the hardware infrastructure and perform optimization both in energy consumption and economic control from a modular architecture which is fully adaptable to most cities. Machine-to-machine (M2M) permits connecting all the sensors of the city so that they provide the platform with a perfect perspective of the global city status. To carry out this optimization, the platform offers the developers a software that operates on the hardware infrastructure and merges various techniques of artificial intelligence (AI) and statistics, such as artificial neural networks (ANN), multi-agent systems (MAS) or a Service Oriented Approach (SOA), forming an Internet of Services (IoS). Different case studies were tested by using the presented platform, and further development is still underway with additional case studies.

Keywords Smart Cities · Intelligent systems · Machine to Machine · Internet of Services · Big data

1 Introduction

The concept of Smart Cities (SC) is a current trend in technological projects. The balance with natural resources and the environment is responsible for these paradigms, which aim to increase the level of comfort for all citizens and institutions based on sustainable development.

P. Chamoso(✉) · J.F. De Paz · S. Rodríguez
Computer and Automation Department, University of Salamanca, Salamanca, Spain
e-mail: {chamoso,fcofds,srg}@usal.es

J. Bajo
Artificial Intelligence Department, Polytechnic University of Madrid, Madrid, Spain
e-mail: jbajo@fi.upm.es
© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_20

When trying to make a city smarter, most of the efforts address energy areas because this will result in high resources saving which allow resources to be applied in other social areas. Therefore, improving the energy efficiency in cities is not only aims to reduce energy costs, but also to promote environmental and economic sustainability.

One of the main problems when developing hardware or software to increase citywide energy efficiency is the integration and deployment in real cities of a developed system, which can be tested and evaluated. In order to facilitate this integration task, a simulation environment that can test and evaluate both the hardware and the developed software or algorithms is presented.

SC can be represented as a three-level platform, where hardware represents the base, the communication mechanism represents the middle layer, and the top layer is represented by the intelligent software [1].

The presented platform consists of a series of heterogeneous sensors that use M2M for the communication, and applies different AI and statistics technologies to evaluate the system and increase energy efficiency in cities.

One of the main characteristics of the simulation environment is that sensors can be easily integrated with the system. In fact, the most commonly used sensors have already been integrated and developers can use them to get the information. In addition, there are some open data services available to provide useful information that sensors cannot provide under simulated conditions, such as the city's weather forecast or traffic.

M2M communication platforms take the role of controlling the communication among all the connected elements of the cities. The decline in costs of connectivity and the price of required devices are driving the growth of this kind of communications. This platform is shaped by combining services to add some intelligence to these connected elements in an interoperable manner that provides the system with a highly heterogeneous content. That intelligence is added in the upper layer of the system to provide the best results depending on the environment status.

The rest of this paper is structured as follows: Section 2 shows the state of the art concerning projects and research conducted in the field of SC, showing the most commonly used techniques and technologies in this field, and then carrying out a comparison between them and the system presented. Section 3 presents the developed system, its operation and details of the techniques used. Section 4 describes the case studies developed to test the platform. Finally, Section 5 presents some results and conclusions of the work.

2 Background

The concept of SC or smart environments [5] itself is still emerging. Making a city "smarter" is one of the main objectives of the researches as a strategy to reduce some problems caused by the rapid growth of the urban population. Problems such as pollution (visible in bigger cities), lack of resources, traffic congestion and deteriorating infrastructure are some of the many problems that large urban populations are increasingly facing [4].

There are multiple definitions of SC, for example as that proposed in [7]: "*the use of smart computing technologies to make city services more intelligent, interconnected and efficient - which includes administration, education, health care, public safety, real estate, transportation and utilities.*" From which we can deduce that new technologies (ICT, Information and Communication Technologies) are the base to provide better quality of life and prudent management of natural resources through the engagement of all citizens.

A smart city is also considered a "system of systems" [6], where different integrated systems form a closed loop and are characterized by functions: sensing, information management, analytics, modeling, and influencing outcomes. Every system produces its own information and consumes the information of others in a well- defined urban planning.

More and more cities around the world are committed to developing pilot projects related to this movement, such as the SmartSantander[1] project in Santander, where there is a large display of parking sensors that indicate available outdoor parking spaces to drivers. They have also deployed a municipal Wi-Fi network that aims to cover the entire city and obtain useful information about the number of Wi-Fi connections from all the parts of the city. Augmented reality applications focused on tourism are also offered. Málaga Smart City[2] is a project that aims to save energy by applying micro power management, which consists of storing energy in batteries for use in buildings, street lighting and electrical transport. Another project is "Smart City Valladolid-Palencia"[3]. As the project is applied to two cities, and not one, it involves the additional task of addressing the issue of transport between the cities. The project has smart a meter network, integration of electric cars, energy efficiency in buildings, traffic organization, etc.

As previously mentioned, SC are typically composed of three levels: hardware, communication and intelligence. While this research work comprises all of them, its main focus is the first level, which addresses infrastructure and the way the information is gathered, but also includes some aspects analysis. The present study introduces a simulation environment that incorporates different technologies, techniques or algorithms and can be easily applied to any city.

With this platform, all innovations can be tested before deploying the infrastructure in a real city, which makes it possible to reduce the budget for the testing phase, allowing all cities to have their innovation impact evaluated and tested the with the platform before being deployed in the real city. The only requests to use the platform are: i) configure different settings about the simulated structures, ii) integrate the used sensor network following the platform rules, and

1 http://www.smartsantander.eu/

2 http://www.lacatedralonline.es/innova/system/Document/attachments/12351/original/IDCCiudadesinteligentes.pdf

3 http://www.valladolidadelante.es/lang/modulo/?refbol=adelante-futuro&refsec=smart-city-vyp&idarticulo=79302

iii) select different on-line open data services that will provide all the information that the algorithms require as input. The system offers the use of previously validated algorithms and simulation models to improve the energy efficiency: however, developers can define those to be used, and then compare them to choose the best option to use in the real city.

Regarding the hardware used for the simulation environment, the combined application of M2M along with a data preprocessing step help to consider the city as a homogeneous network of sensors, even though the sensors used are heterogeneous [2][3]. The platform offers the use of different commonly used sensors to gather the most relevant information of any city, but new sensors can be integrated easily.

The platform has been tested and validated successfully and continues its development today. The following sections describe the operation and technologies used in the platform, and the results currently obtained in different case studies.

3 Proposed Architecture

The proposed architecture is structured by following the previously identified three layer model as shown in Fig. 1. To begin, the base of the architecture is represented by the city infrastructure, which is composed of every different objects of the city to be monitored in conjunction with the required sensor and the technology used, in order to exchange the data with the system.

Functionalities that are used to define and configure the sensor networks integrated in the objects are located in the middle layer. Due to the multiple kinds of objects that compose the network, communications are heterogeneous in nature. Consequently, it is necessary to pre-process the data to allow the system to use data independently of the technology used for the transmission, and thus allowing the system to be appreciated as a single and homogeneous network.

The third and final layer is the associated to the intelligence of the city, where all the information is analyzed and managed in such a way that the highest benefits are provided and the most appropriate services are offered to the citizens.

The following section describes all the layers and their sub-systems.

3.1 Infrastructure Layer

Starting from the lowest layer, the first level corresponds to both the connected objects of the city and the sensor networks.

The number of objects that can be connected to the network to be monitored as part of a global system is invaluable. Not only the number, but also the type of these connected objects grows substantially every day. The most common connected objects of smart cities were included in the developed simulation environment.

This simulation environment consists of a physical part, connected to a software part, so it differs significantly from the simulation systems used to date, which only include software simulations. Thus, with this kind of environment, multiple real sensors can be included and users can interact with the simulation environment, getting more realistic simulations.

The physical part is composed of different kinds of sensors and actuators and a city mockup, while all its elements were created with a 3D printer simulating real objects, buildings and cars. All these elements are connected in several different ways and they provide both real and simulated values. Real values are provided by the developers, who integrate their own prototypes with the mockup.

In addition, for every kind of element, the kind of technology that is more frequently used to gather or exchange information was analyzed and included as part of of the system. This ensures that widely applied technologies and techniques can be easily tested with the developed simulation environment.

The smart city common objects that have already been embedded in the simulation environment are:

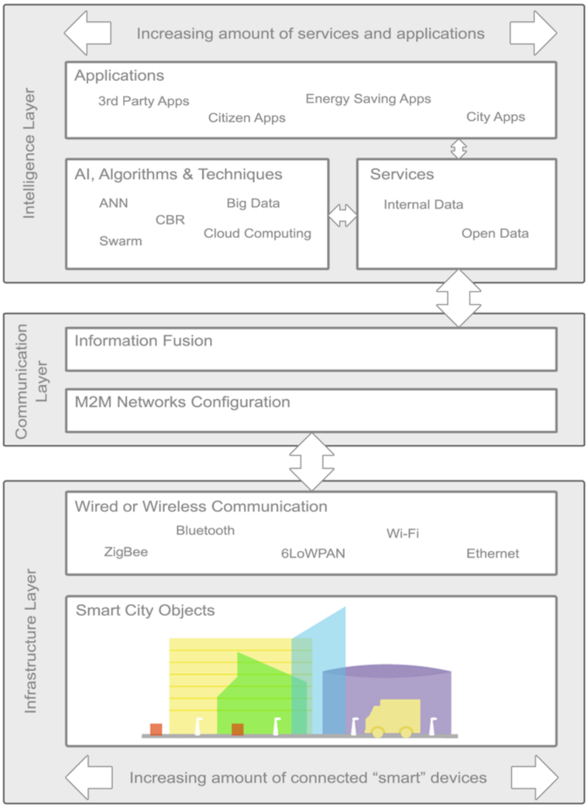


Fig. 1 System architecture overview

Buildings: There are different kind of buildings distributed among the city mockup, each of which is organized in different districts according to its category. For example, there are residential, business, commercial, industrial or energetic districts. Each has its own features, for example when consuming or generating energy. These values can be simulated or provided by the developers when testing the system. The simulated values are based on different parameters like the day, the current time or the configured energetic infrastructures.

Public Lighting: A series of lamps was distributed along the streets of the simulation environment. The system makes it possible to both simulate and analyze the public lighting to detect failures and find optimal solutions. Moreover, the system allows the testing of different energy distribution algorithm according to different parameters, such as the assigned budget or traffic.

Containers: There are several groups of recycling and waste containers distributed among the city districts. Their value can be simulated according to the consumption habits of every district or even directly associated to real sensors that measure the fill level of the containers. Different calculations can be evaluated by using those data. For example, the route of every track can be optimized to collect only the waste of the full containers, or to change the location of those containers that are always empty.

Energy Generation Elements: Several elements that simulate the production of electric power from the sun or the wind were included in the simulation environment. Data are simulated by using real open data services of the city that it is being simulated.

Vehicles: One of the most important aspects to bear in mind when trying to make a city smarter is its traffic management and distribution. Vehicles, whether public or not, play an important role when evaluating the quality of the lives of the citizens and that of the environmental, which is often harmed due to high pollution rates. Accordingly, all the vehicles included in the simulation environment are monitored and controlled to carry out different simulations. In order to guide them among the roads of the city, a steel rail is embedded in a methacrylate base and covered by a vinyl that simulates the road. Junctions are possible as a result of the servomotors that move the rail to follow the desired direction. Front wheels are guided by a small magnet that follows the steel rail, and rear wheels are directly connected to a DC motor whose speed can be controlled by software. To know the position of every car, several magnets were embedded in parallel to the steel rail, forming binary patterns that identify the street. To detect those magnets, cars are equipped with a hall sensor, allowing their position to be known at all times.

The global sensor network may be composed of multiple heterogeneous sets of individual sensors associated to one or more objects of the city. These sensors are in charge of automatically gathering the requested data of the nearest environment by themselves.

Table 1 identifies the sensors that were adapted to the main common objects, and the communication methods used in the test.

Table 1 Mockup connected elements, built-in sensors and the communication type.

Elements	Sensors	Communications
Buildings	Ammeter, temperature, humidity	Serial*, Wi-Fi, Ethernet
Lamps	Ammeter, voltmeter, LDR, CO_2	PWM*, Wi-Fi, Ethernet
Containers	Ultrasonic, temperature	Wi-Fi, BLE, 4G, GSM
Solar panels	Ammeter, voltmeter	4G, GSM
Vehicles	Hall*, voltmeter, ammeter	BLE, Wi-Fi

Sensors and communications marked with an asterisk (*) are only used for their integration with the simulation environment.

3.2 Communication Layer

Sensors must also incorporate some kind of mechanism to transmit this information so that it can be integrated in the global system for future use. On many occasions, the cities already use sensor networks that have been incorporated in order to develop studies or to take measures. They could also have been added due to some other reason, prior to the process of transforming the city into a smart city. In addition to the previously commented evolution in technology, this could result in the cohabitation of different kinds of sensors, gathering the information associated to a same environment.

Large-scale systems typically include middleware (MW) to collect information coming from the sensors. The main task of this MW is to allow the information gathered by the sensors to be standardized within the system. To do so, the system includes an intermediate layer specialized in collecting information by using the required protocol for each technology. Once the information is obtained by the sensor and transmitted, the MW collects it and processes it (without interruption) to produce a new encapsulation to a common protocol for the whole system.

This way, the inclusion of new technologies, which may appear in the future to offer new or better features, prevents the bulk of the smart-cities management system from being affected in the case of using them. In such cases, the procedure would be based, as a general rule, on creating a new module incorporating the MW as a driver. This module will include the programming associated to the conversion of information formats and communication protocols. In turn, it will allow the definition of every configuration parameter that such technology requires.

The inclusion of an MW with such characteristics is necessary in the system, as it means a reduction in risk of some factors that affect the conflict of decision making regarding the technology to be used in a city. In particular, the communication protocol directly affects important factors such as sustainability, adaptability and system maintenance over time.

Hosting and information processing services are performed in a Cloud Computing (CC) environment; therefore, such environments support the communication and subsequent levels. The application of CC technology, particularly to smart city management, provides a deployed system with different benefits. To begin, the initial investment will be reduced due to the fact that there is no need to acquire hardware and software to deploy the system. Flexibility is the other main advantage provided by CC because the service is capable of adapting itself to the demand and always providing the necessary resources. As a result, the size of the city is no longer relevant.

Using a practical case taken from this work as an example, the MW is divided into a series of modules, each of which is associated to a different technology. For example, to communicate through Bluetooth 4.0 (BLE - Bluetooth Low Energy), there is a module in charge of requiring the technology-specific information needed to register a sensor of this kind. For the case of BLE, this information only includes the MAC (Media Access Control) address. Once the configuration has been stored, the *Gatttool* library is internally used. This library makes it possible to use any BLE device through a series of simple commands. Similarly, there are specific modules to communicate through ZigBee, RFID or Wi-Fi technology, as well as through wired sensors, independently of their protocol or API. In every particular case, the specific de-encapsulation, associated to the technology, is performed. To do so, the obligatory information is extracted and the encapsulation to the protocol required by the higher layers in the system is performed. These processes are all performed remotely (in the CC environment), which allows us to take advantage of the benefits each technology offers regarding consumption, range or duration, without affecting the gathering of data by the higher levels of the architecture.

Thus, it is necessary to include a MW with such features in the system, as it implies the resolution of the protocol incompatibility conflict, which directly affects the stability of the entire system. However, despite using these techniques, additional conflicts still remain.

3.3 Intelligence Layer

The last layer is composed of three big blocks: i) the set of technics and algorithms used for the data treatment and processing; ii) the generation of several services based on the platform data or the integration of third party services to allow real simulations; iii) the set of final applications.

First, the set of technics and algorithms implemented in the platform is scalable and reusable. So far, there are different technics and algorithms that allow developers to store, manage and analyze all the system data. As the number of connected objects can grow, the storage infrastructure and the data treatment algorithms must be ready to grow at the same pace. The last layer is deployed in a CC platform and uses mechanisms to deal with big data.

Developers that use the platform can specify their own AI algorithms or use those that have been proposed to test them or compare their effectiveness.

To use the proposed algorithms, the platform architecture must be based on SOA. A set of internal and third party data based services are offered, which the developers are free to access because of the IoS.

These services are used by all the developed applications, which increase the quality of services delivered to the citizen or to the city in general. For example, the simulation environment provides different apps such as citizen recommendation, citizen participation, energy management, efficiency measurement or public services path planning.

4 Results, Conclusions and Future Work

Different vertical applications have been developed and integrated with the simulation environment. This article is focused on the development of a platform that can optimize various aspects of city management, including the collection routes used by waste trucks. In this case, the optimization is performed by taking into consideration the fill level of the glass, paper, plastic and waste containers of the city.

This section details the case study, specifically the hardware integration with the simulation environment platform, which is the lower layer, as the upper layer is designed for the applications of other developers. Finally, conclusions and future work are presented.

4.1 Waste Collection Optimization

As previously mentioned, the mockup is composed of different districts with different containers: paper, glass, plastic and waste. During simulation, they are filled in a pseudo-random way, taking into account different parameters associated to their district, the time or the weekday of the simulation.

Two different kinds of sensors were integrated, both of which are associated to the value of two mockup containers. Different tests to place the sensor in a real container were evaluated, as well as the type of communication.

Figure 2 shows the schema of the sensor used for the integration with the two evaluated communication types: i) by using a GSM 3G module, which requires the use of a microcontroller to analyze the measured data, and ii) a Wi-Fi module with integrated processing capacity, which is capable of reading the measured value, processing and transmitting it only when there is a change, which saves the battery life.

The sensor used in this case study is an HC-SR04, which works at a frequency of 40kHz in a range from 1.7cm to 4.5m, thus making it very suitable for use in containers.

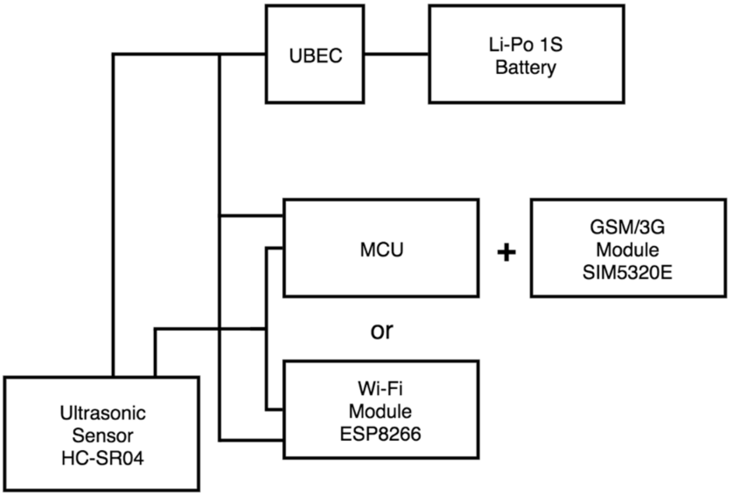


Fig. 2 Waste Level Sensor Schema.

In order to consider a measured value as different from the previous one, the difference must be more than 2 centimeters and the change must last a minimum of 10 seconds. This results in a considerable saving of energy and the tested battery, with a capacity of 3200mAh can last up to 26 days.

Over this infrastructure and communication layers, different algorithms for optimizing the path taken by the waste trucks were been evaluated. As the positions of the trucks are also known by the system, only those containers with a level higher than 50% are included in the route.

4.2 Conclusions and Future Work

This physical simulation environment of smart cities, completely integrated with a software platform, is unlike other simulation systems that only offer software simulations.

The main advantage of this work is the possibility of including physical simulation tests, which are more complete than mere software simulations.

Given the presented case, in which volumetric sensors were placed inside the containers of the city, the company responsible for the recycling would benif from cost savings with both personnel and resources, as only one set of containers would be included in the truck route.

As future work, a virtual organization based multi-agent system will be included in order to organize the platform autonomously in line with the requirements of the city's evolution, regardless of the city characteristics or values.

Acknowledgements This work has been supported by the European Commission H2020 MSCA-RISE-2014: Marie Skłodowska-Curie project DREAM-GO Enabling Demand Response for short and real-time Efficient And Market Based Smart Grid Operation - An intelligent and real-time simulation approach ref 641794.

The research of Pablo Chamoso has been financed by the Regional Ministry of Education in Castilla y León and the European Social Fund (Operational Programme 2014-2020 for Castilla y León, EDU/310/2015 BOCYL).

References

1. Calvillo, C.F., Sánchez-Miralles, A., Villar, J.: Energy management and planning in smart cities. Renewable and Sustainable Energy Reviews **55**, 273–287 (2016)
2. Chamoso, P., De la Prieta, F., De Paz, F., Corchado, J.M.: Swarm agent-based architecture suitable for internet of things and smartcities. In: 12th International Conference on Distributed Computing and Artificial Intelligence, pp. 21–29. Springer International Publishing (2015)
3. Chamoso, P., De la Prieta, F., Bajo, J., Corchado, J.M.: Conflict Resolution with Agents in Smart Cities. Interdisciplinary Perspectives on Contemporary Conflict Resolution. IGI Global (2016)
4. Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J.R., Mellouli, S., Nahon, K., Pardo, T.A., Scholl, H.J.: Understanding smart cities: an integrative framework. In: 2012 45th Hawaii International Conference on System Sciences, pp. 2289–2297 (2012)
5. Gómez-Romero, J., Serrano, M.A., Patricio, M.A., García, J., Molina, J.M.: Context-based scene recognition from visual data in smart homes: an information fusion approach. Personal and Ubiquitous Computing **16**(7), 835–857 (2012)
6. Naphade, M., Banavar, G., Harrison, C., Paraszczak, J., Morris, R.: Smarter cities and their innovation challenges. Computer **44**(6), 32–39 (2011)
7. Washburn, D., Sindhu, U., Balaouras, S., Dines, R.A., Hayes, N., Nelson, L.E.: Helping CIOs Understand "Smart City" Initiatives. Growth **17** (2009). http://c3328005.r5.cf0.rackcdn.com/73efa931-0fac-4e28-ae77-8e58ebf74aa6.pdf

Hash-Chain Based Authentication for IoT Devices and REST Web-Services

António Pinto and Ricardo Costa

Abstract The number of everyday interconnected devices continues to increase and constitute the Internet of Things (IoT). Things are small computers equipped with sensors and wireless communications capabilities that are driven by energy constraints, since they use batteries and may be required to operate over long periods of time. The majority of these devices perform data collection. The collected data is stored on-line using web-services that, sometimes, operate without any special considerations regarding security and privacy. The current work proposes a modified hash-chain authentication mechanism that, with the help of a smart-phone, can authenticate each interaction of the devices with a REST web-service using One Time Passwords (OTP). Moreover, the proposed authentication mechanism adheres to the stateless, HTTP-like behavior expected of REST web-services, even allowing the caching of server authentication replies within a predefined time window. No other known web-service authentication mechanism operates in such manner.

1 Introduction

The Internet of Things (IoT) can be seen as a distributed network of devices that interact with human beings and with other devices [1, 2]. New applications for such devices appear on a daily basis and these, typically, use sensors to collect data. The IoT is expected to become a key source of big data and analytics [3]. Example sensors

A. Pinto(✉)
GCC, CIICESI, Escola Superior de Tecnologia e Gestão de Felgueiras,
Politécnico do Porto and INESC TEC, Porto, Portugal
e-mail: apinto@inesctec.pt

R. Costa
GCC, CIICESI, Escola Superior de Tecnologia e Gestão de Felgueiras,
Politécnico do Porto, Porto, Portugal
e-mail: rcosta@estgf.ipp.pt

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_21

are accelerometers, gyroscopes, magnetometers, barometric pressure sensors, ambient temperature sensors, heart rate monitors, skin temperature sensors, GPS, video cameras, microphones, among others.

A possible classification of IoT devices can be done with respect to their communication capabilities. The adopted reference scenario is depicted in Figure 1 and comprises three types of sensors. Type A sensors are characterized by requiring a specific Wireless Sensor Network (WSN) gateway, typically from the same manufacturer of the devices, and by being built for ultra low power operation. These run on (coin shaped) batteries and minimize wireless communications in order to expand their lifetime. The security, authentication and confidentiality of the collected data is achieved by means of pre-built, per device, cryptographic encryption keys that are exchanged with the WSN gateway upon initial set-up. Type B are characterized by being more autonomous, not requiring a gateway, and by being able to interact directly with the on-line central server. These are either connected to a power outlet or run on batteries with larger capacity, which are also recharged more frequently (daily or more). The security, authentication and confidentiality of the collected data can be achieved by any available mechanism. Type C are characterized by requiring a type B device in order to upload the collected data to the on-line central server. These use short range wireless communication capabilities, such as Bluetooth, to communicate with a type B device that has standard IP connectivity. The security, authentication and confidentiality of the collected data can be achieved by any mechanism available in the type B device.

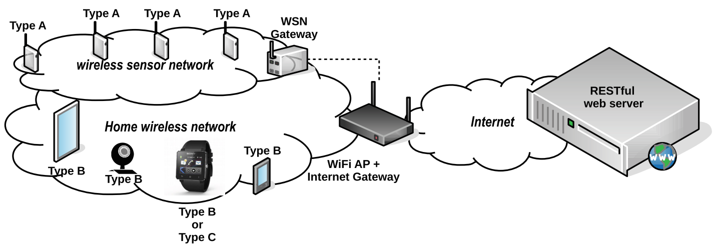


Fig. 1 Adopted reference scenario

The data collected by all these sensors tend to be personal, sensible and private. The privacy and control over the collected data was already addressed by the authors with the Sec4IoT framework [4]. Nevertheless, the need to assure that the collected data is always maintained within the control scope of the Sec4IoT framework, an end-to-end device authentication mechanism is required.

The paper is organized in sections. Section 2 describes and compares the work of others that is related to ours. Section 3 describes the proposed authentication mechanism and performs its security validation. Section 4 describes the experimental set-up, the implemented authentication prototype and presents the obtained results. Section 5 concludes the paper.

2 Related Work

REST is a distinct way of deploying web-services that is becoming quite common due to its simplistic and HTTP-like behavior. The key REST principles are threefold: 1) explicit use of HTTP methods; 2) stateless; and 3) resources must be named using URLs (Uniform Resource Locators). The hyped stateless operation of a REST server may not be completely possible, especially if one considers clients authentication and authenticated sessions management. Current solutions make use of standardized HTTP related authentication mechanisms, such as Basic and digest access authentication [5], or of Open standard for authentication (OAuth) [6, 7], or use proprietary, in-house developed, solutions.

The HTTP protocol supports several authentication mechanisms [8] in order to control the access to pages and other resources. This solutions make use of the 401 status code and the WWW-Authenticate response header [5]. In form-based authentication, developers, instead of relying on authentication at the protocol level, can make use of web-based applications or HTML code embedded into their web pages. They can use INPUT elements in HTML Forms to request the client's credentials (User and Password) as a normal part of their web application. The Open standard for Authentication (OAuth) [6, 7, 9] provides a method for a web application to grant third-party access to their resources without, actually, sharing their clients credentials. In [10] the authors propose the use of a token based approach for authentication in REST-based web services. Their proposal consists in extending the HTTP authentication to include a UsernameToken as a secondary password verification. This would allow providers to customize their own authentication according to their specific need, improving flexibility and security, and introducing the possibility of the server challenging the client in order to authenticate it. The key advantage of this solution was lost as it addresses the same shortcomings of the first HTTP Digest authentication, as does HTTP1.1.

Table 1 compares the solutions identified as related to ours. For instance, the version 1.1 of the HTTP digest authentication performs data encryption (2^{nd} column), authenticates both server and client (3^{rd} column), resists attacks that resend previously exchanged messages (4^{th} column), is secure against eavesdropping and Man in The Middle (MiTM) attacks (5^{th} column), does not require trust third parties (6^{th} column), the authentication is performed at the HTTP layer (7^{th} column) and

Table 1 Related work comparison

Solution	Data enc.	Mutual auth.	Replay resistant	MiTM resistant	3rd parties	Auth. control	OTP
Basic	N	N	N	N	N	HTTP	N
Digest	Y	N	N	N	N	HTTP	N
Digest 1.1	Y	Y	Y	Y	N	HTTP	N
OAuth	Y	N	Y	Y	Y	App	N
AuthToken	Y	N	Y	Y	N	HTTP	N

has no support for OTP (8^{th} column). Regarding the authentication control, it can either be controlled by the web server (identified in the table with HTTP) or by the web-service (identified with App). None of the presented solutions supports transaction control by means of OTP. The authors believe that such OTP per transaction approach is the one that better suites the REST design philosophy.

3 Minimalist Authentication Mechanism

The proposed Minimalist Authentication Mechanism (MAM) requires a secure client register procedure, deployed as a secure web page (HTTPS), that is assumed to be in operation. The secure register procedure will enable the secure generation and exchange of a per device secret (A_{sec}). The proposed algorithm implies that any request made by clients to the server must comprise, aside other parameters, the client identification and an OTP. A, per device, set of OTPs is generated with the login procedure. The login is initiated when the client calls the login procedure made available by means of a REST-based web-service. The client computes the non-guessable random value $nonce_{A1}$, calculates both the cli_n_2 and the $Time$ values and passes them as parameters to the login procedure. The nonce generation function is assumed to be secure. $Time$ value is obtained by rounding up the current time in intervals of 10 minutes. The server will compute a local n_2 value using a secure hash function over $nonce_{A1}$, n_1 and the $Time$ value. The computed n_2 value, if equal to the cli_n_2, will be used to create the initial security token (tk_0). After the generation of the initial token, tokens tk_1 to tk_{512} will be calculated by using the cryptographic hash function over the previous token. Both server and client will store the set of 512 tokens to be used as OTPs, in reverse order, in the subsequent 511 requests of that client. Such will enable anti-replay protection and prevent both man-in-the-middle and Denial of Service (DoS) attacks. Additionally, and due to the fact that the server returns both the $seed^{Time}$ and the token tk_{512}^{Time} to the client, both server and client are mutually authenticated.

The way the OTPs are obtained depends on the type of sensor (Figure 1). On the one hand, type B sensors have the required capabilities, in terms of wireless communications, CPU and available battery lifetime, to perform a complete login by themselves. At the end of the login procedure, the device will have a set of OTPs to be used in the subsequent requests to the REST web-service. On the other hand, both Type A and C require additional devices in order to complete a successful login procedure. Currently, type A devices require a specific WSN gateway, whereas type C devices require a paired smart-phone.

In the proposed solution, a smart-phone will be used to complete a successful login and to obtain a set of OTP which will then be sent to the type A or C sensor by any means available in the sensor. Such will avoid two major drawbacks of the current solutions. Firstly, the WSN specific gateway will no longer be required as it can be replaced by a existing wireless Access Point (AP) with Internet connectivity. Secondly, none of the cryptographic material to be stored on the sensors, the set of

OTPs, can be used to perform a new login in the system. A sensor login procedure that will authenticate the smart-phone is assumed to be in operation. Nevertheless, the REST web-service does not take part in the sensor/smart-phone authentication procedure.

Type A sensors run on ultra low power hardware, use small sized batteries and communicate periodically in order to save power. The size of the OTPs set must be adapted to each case. For instance, if a sensor communicates with the server once per hour, it will require 24 OTPs per day, 167 per week, or 672 OTPs per month. The REST web-service will have a set of URLs that will enable the login procedure to reply with sets of OTPs of different sizes and using different secure hash functions.

3.1 Security Analysis

The Automated Validation of Internet Security Protocols (AVISPA) [11] tool was used to perform the security validation of the proposed authentication mechanism. The AVISPA tool performs the automated validation of security protocols described in High Level Protocol Specification Language (HLPSL) [12]. HLPSL enables the description of both the protocol and the required security properties, such as secrecy and authentication. AVISPA adopts the Dolev-Yao intruder model [13] where the intruder is in complete control of the network. Our HLPSL specification is available online[1].

The client-server communication confidentiality is obtained by means of a pre-shared secret between that specific client and the server. This pre-shared secret is assumed to be available on the client and on the server and to be refreshed frequently. For instance, such pre-shared key may be refreshed upon every client successful login. If an attacker obtains a capture of the exchanged messages, and while being able to obtain the $nonce_{A1}$ and cli_n_2 values, these are the result of one way hash functions. Meaning that eavesdropping attacks are not possible.

An attack such as a man-in-the-middle attack is only possible if the pre-shared key is compromised. The proposed solution assumes that this key is secure. Nonetheless, the pre-shared key is never transmitted on the link an is assumed to be a result of a specific pre-shared key creation procedure that may be triggered by the user whenever he wants to, by means of a user management web site requiring a two-factor authentication.

Despite the fact that both sides make use of random values to generate or verify an authentication token (tk_{512}), these values are never exchanged in clear text. The only value exchanged between client and server, besides the authentication token, is the seed value. The seed value is, in turn, the result of a secure hash function. Neither entity can force the other in to generating a specific seed or authentication token. Meaning that attacks that are based in key control are also not possible.

---

[1] Available at http://www.estgf.ipp.pt/~apinto/mawr.hlpsl

Table 2 Request processing capabilities by the server

Digest algorithm	Requests/sec	Request (ms)
SHA-256	207.5	4.7
SHA-512	211.3	4.8

The proposed solution makes use of the the $nonce_{A1}$ that is assumed to be securely generated and only used once. If a second request is received with a repeated *nonce* value, the server will ignore the request and, thus, reject replay attacks.

DoS attacks are based on overwhelming a server with requests so that it is not able to respond to legitimate requests. The proposed solution, uses the $Time$ value, rounded up to 10 minute intervals, so that all requests sent by the same client within this time interval (up to 10 minutes) will have the exact same reply. Due to being a REST-based solution, it can easily be deployed within a Content Delivery Network (CDN) [14], i.e. such reply could be cached, making it very difficult to perform a successful DoS attack. This approach limits the number of per device successful logins to one authentication per time period of 10 minutes. This is a drawback of the proposed solution but the time period can be reduced and fined tuned to each implementation.

4 Experimental Results

The prototype was developed using Java and the *Netbeans Java* IDE. The Web application ARchive (WAR) file was built and deployed on a Glassfish 4.1 server, running on a 64 bit Linux system (kernel 3.18.8-201.fc21.x86_64) with 8GB of memory and a dual-core Intel Pentium G645 2.9GHz processor. All results shown in this section were obtain by running multiple sets of 1000 executions each. SHA-256 and SHA-512 where the selected secure hash algorithms, from the list of the algorithms available in the *Java* language, mainly because the remaining ones are currently considered insecure by multiple sources. Stevens demonstrated a collision attack on the MD5 algorithm [15] and Liang et. al, later presented an improved collision attack to the same algorithm [16]. The SHA-1 algorithm was also demonstrated to be less complex that the initial expectation. In particular, Wang et al. demonstrated that the theoretical number of 2^{80} hash operations that where assumed to be required to find a collision could be reduced to a lesser value of 2^{69} operations [17].

Table 2 shows the average number of requests processed per second and the average time required by the server to process one request, for both the SHA-256 and the SHA-512 secure hash algorithms. As can be seen, a user login procedure takes approximately 5ms to be completed and the server is able to process about 210 requests per second. While there is a difference between the results obtained using different secure hash algorithms, this difference is very small can be neglected.

Table 3 Average request processing time required per HTTP authentication mechanism

HTTP Authentication Mechanism	Request (ms)
Basic	4,8
Digest HTTP1.0	64,7
Digest HTTP1.1	64,0
MAM (SHA-256)	4,7

The bandwidth usage tend to be slightly more (approx. 8% more) when using the SHA-512 secure hash function. If we consider that login procedure, described in the previous section, returns the token tk_{512} that is the result of the used hash function and, evidently, will be larger when using the SHA-512 when compared to the use of SHA-256. The remaining elements of the messages exchanged between server and client are of equal size and independent of the secure hash algorithm that was used.

Table 3 shows the average processing time for each authentication mechanism made available by the HTTP protocol. The results show that both digest authentication mechanisms supported by HTTP take 64 ms, or longer, on average, to process a client authentication. The insecure basic authentication, that does not encrypt user credentials, was the only one that obtained processing times similar to those obtained by the proposed solution.

The results shown in Table 3 were obtained on a system running a 64 bit Linux with 8GB of memory and a dual-core Intel Pentium G645 2.9GHz processor. A HTTP server was setup in a virtual machine running Linux. The Apache web server was installed and configured to support the three authentication mechanisms identified. The Linux *curl* command was used as the HTTP client. All results shown in this section were obtain by running 3 sets of 1000 executions each. The proposed solution (MAM) is the fastest one, requiring only 4,7 ms to conclude. It is even slightly faster that the HTTP basic authentication.

5 Conclusion

The IoT is becoming the next big technological hype. New services that make use of IoT devices appear every day. These new services use web-based storage and the IoT devices are starting to interact directly with such web-storage. Two problems arise from such a scenario: 1) the privacy and control over the collected data; and 2) end-to-end authentication for low specked devices. Our previous solution (Sec4IoT) dealt with the first problem, the later subsisted.

This work addresses the second problem while still maintaining a REST-like API so it can be integrated within any existing web-service. Moreover, the proposed authentication mechanism does not require trust in third-parties, maintains the authentication control in the web application and, by requiring low computational power, it can be deployed in low end IoT devices.

References

1. Xia, F., Yang, L.T., Wang, L., Vinel, A.: Internet of things. International Journal of Communication Systems **25**(9), 1101–1102 (2012)
2. ABIresearch, The internet of things will drive wireless connected devices to 40.9 billion in 2020 (2014)
3. Press, G.: It's Official: The Internet of Things Takes Over Big Data as the Most Hyped Technology (2014)
4. Costa, R., Pinto, A.: A framework for the secure storage of data generated in the iot. Advances in Intelligent and Soft Computing (2015)
5. Leach, P.J., Franks, J., Luotonen, A., Hallam-Baker, P.M., Lawrence, S.D., Hostetler, J.L., Stewart, L.C.: HTTP Authentication: Basic and Digest Access Authentication
6. Hardt, D.: The OAuth 2.0 Authorization Framework
7. Hardt, D., Jones, M.: The OAuth 2.0 Authorization Framework: Bearer Token Usage
8. Fielding, R., Reschke, J.: Hypertext Transfer Protocol (HTTP/1.1): Authentication
9. Jammer-Lahav, E.: The OAuth 1.0 Protocol
10. Peng, D., Li, C., Huo, H.: An extended usernametoken-based approach for REST-style web service security authentication. In: 2nd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2009, pp. 582–586, August 2009
11. Armando, A., Basin, D., Boichut, Y., Chevalier, Y., Compagna, L., Cuellar, J., Drielsma, P.H., Heam, P.C., Kouchnarenko, O., Mantovani, J.: The Avispa Tool for the Automated Validation of Internet Security Protocols and Applications, vol. 5, pp. 281–285. Springer (2005)
12. Chevalier, Y., Compagna, L., Cuellar, J., Drielsma, P.H., Mantovani, J., Modersheim, S., Vigneron, L.: A high level protocol specification language for industrial security-sensitive protocols. In: Proc. SAPS, vol. 4, pp. 193–205
13. Dolev, D., Yao, A.: On the security of public key protocols. IEEE Transactions on Information Theory **29**, 198–208 (1983)
14. Pathan, M., Buyya, R., Vakali, A.: Content delivery networks: state of the art, insights, and imperatives. In: Buyya, R., Pathan, M., Vakali, A. (eds.) Content Delivery Networks. Lecture Notes Electrical Engineering, vol. 9, pp. 3–32. Springer, Heidelberg (2008)
15. Stevens, M.M.J.: Fast Collision Attack on MD5, tech. rep., March 2006
16. Liang, J., Lai, X.-J.: Improved Collision Attack on Hash Function MD5. Journal of Computer Science and Technology **22**, 79–87 (2007)
17. Wang, X., Yin, Y.L., Yu, H.: Finding collisions in the full SHA-1. In: Shoup, V. (ed.) Advances in Cryptology–CRYPTO 2005. Lecture Notes in Computer Science, vol. 3621, pp. 17–36. Springer, Heidelberg (2005)

Smart Computer-Assisted Cognitive Rehabilitation for the Ageing Population

Miguel Oliver, Pascual González, Francisco Montero, José Pascual Molina and Antonio Fernández-Caballero

Abstract This article introduces some solutions based on assistive technologies to help ageing populations in relation to their inherent problem of interaction. These technologies envision physical and cognitive rehabilitation. Examples of research and commercial systems in both types of rehabilitation are shown in this article. Finally, and based on our prior experience, a recently developed system for cognitive rehabilitation with physical exercise is presented. The main characteristics of this system are that it adds intelligence provided by a Rule Authoring component, and it offers the possibility of expansion in the future due to its implementation.

Keywords Gerontechnology · Assistive technologies · Cognitive rehabilitation · Physical exercise

1 Introduction

The interest of society to live longer and better is always growing. Formerly, people wanted to live longer. However, with increasing life expectancy the interest in living a better old age has also grown. In addition, in the field of computing there is a high interest in working on several aspects to make people live better. One of this research highlights is related to technology associated with the older people (gerontechnology).

M. Oliver(✉) · P. González · F. Montero · J.P. Molina · A. Fernández-Caballero
Instituto de Investigación en Informática de Albacete, Universidad de Castilla-La Mancha, 02071 Albacete, Spain
e-mail: oliver@dsi.uclm.es

P. González · F. Montero · J.P. Molina · A. Fernández-Caballero
Departamento de Sistemas Informáticos, Universidad de Castilla-La Mancha, 02071 Albacete, Spain

© Springer International Publishing Switzerland 2016

H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_22

The principal mission of assistive technologies is to bring technology to people presenting a problem of interaction with the computer. More specifically, assistive technologies for rehabilitation attempt the rehabilitation process to be supported by a computer system. The rehabilitation process can be physical (affecting the movement of bones and joints) or cognitive (affecting brain processing). This paper presents a system to aid in cognitive rehabilitation of older people. The main advantage of this system, over the existing ones, is that it provides cognitive rehabilitation with physical exercise. In addition, this software has an architecture that provides intelligence to help the user when performing the rehabilitation, and the therapist during monitoring of the rehabilitation. It also has an architecture that allows its expansion in the future through the inclusion of new types of exercise.

2 Assistive Technologies in Rehabilitation

Assistive technologies attempt to bring technology to people who can not access it due to physical or cognitive impairments. In the case of computer-assisted rehabilitation systems, the intention is to improve the physical or cognitive condition of the patients through the use of expert systems that help specialists and patients to achieve effective rehabilitation.

2.1 Assistive Technologies in Physical Rehabilitation

There are multiple systems for the physical rehabilitation of patients. These systems attempt patients to regain the lost physical mobility, either due to the inherent deterioration with age, or as a consquence of surgery. People who are born with a disability belong to another group of patients who might benefit from the system. In this case, the system's mission is that the patients acquire a mobility as functional as possible.

There are works that use Kinect depth sensors for the study and development of computer-based rehabilitation systems (e.g., [1], [2], [3]). In addition, there are also commercial systems such as those proposed by KineLabs, Vera, Toyra, TeKi or VirtualRehab.

2.2 Assistive Technologies in Cognitive Rehabilitation

In the event that a patient has a cognitive impairment, there are also systems that help in the rehabilitation process. Users who make use of these systems are people who need to recover lost cognitive abilities. This may be due, as in the previous case, by the degradation resulting from age, or from an injury that affects their cognitive abilities. Those who suffer cognitive deficits since birth can also make use of these systems. In this case, the system's mission is to correct deficiencies. It is impossible to define a unique way of dealing with these problems. This is due to the multitude of cognitive deficits and shortcomings that occur in the human brain, as well as the variety of ways in which they are presented. Each type of brain disorder requires different methods and specialists for their treatment [4].

Among the systems that can be used for the treatment of cognitive problems, we emphasize the software for neuropsychological rehabilitation of cognitive impairment named Gradior [10], that is intended for people with traumatic brain injury, schizophrenia or dementia. The JClic system [11] is another example oriented towards primary education patients. Meanwhile, Bungalow Software [12] and Learning Fundamentals [13] are software programs for the rehabilitation of language. One of the most complete systems is COGREHAB [14] which collects 50 cognitive rehabilitation programs. Also noteworthy is the Tango: H system [5] that makes use of a Kinect sensor to recognize the user's body. This program focuses on hospitalized children with disabilities. In fact, cognitive rehabilitation is a field in which continuous research is conducted. A recent example is a multisensory treatment for people with space hemineglect [6]. Another example is the HABITAT system [4], developed to assist people with acquired brain injury (ABI).

Therefore, in this paper we present a system that combines cognitive and physical rehabilitation. Connecting these two aspects of the rehabilitation process enrich this type of systems, offering the therapist the possibility of designing more complex therapies that improve the rehabilitation of a specific patient. In addition, for the physical rehabilitation therapies, the inclusion of some cognitive tasks provides patients with a more enjoyable environment, where the physical rehabilitation becomes a more game-based experience. Finally, our proposal allows the therapist to design therapies that include the definition of more complex exercises since they can control different difficulty and precision levels, and the decision of increasing or decreasing these levels is made by controlling the specific variables of the exercise and vital signs of the patient.

3 Proposed Cognitive Rehabilitation System

The proposed system combine cognitive rehabilitation with physical exercise. In addition to exercising the brain, the user also exercises the body for complete rehabilitation. The rehabilitation specialist creates each therapy, based on existing literature and the particular nature of each older people. We have used the same therapy decomposition as in [10]. A Therapy is defined as a set of Activities to rehabilitate a specific physical or cognitive impairment. Each Activity can be divided into elemental Tasks. Each Task is a sequence of Steps, which are user's gestures or postures that can be either controlled by the Kinect or can be only informative. In our case we associate the concept Exercise to the Step, and we include a new level of detail, named Sub-step that allows controlling a specific exercise.

Moreover, our system allows the design of rules to control the therapy execution. These rules are the knowledge base of the system to decide, at runtime, which is the best transition between the available ones, i.e. which is the most desirable next action that a user should make while he/she is performing a Step, a Task or an Activity. Each exercise consists in a series of visual objects that have a specific meaning, which the ageing adult should recognize and carry out a specific cognitive process such as association, categorization, and so on. For this, the therapist selects the size, position, orientation and colour of each of the items on the screen.

For its part, the user is placed in front of a Kinect sensor that captures his/her shape and displays it on the screen. The elements defined by the therapist, along with the instructions that guide him/her in developing the exercise appear around him/her. To select items, he/she places a particular body part on an item. The patient's mission is to select all items on the screen following an order imposed by the therapist.

3.1 Types of Exercises

Currently, the system allows developing three types of exercises: Pair Association Exercises, Multiple Association Exercises, and Categorisation Exercises. The exercises have been selected because they are widely applied in the rehabilitation of many cognitive deficits of front executive function of the brain [4]. In addition, it offers a kind of free exercise, in which the therapist can define all options of all elements.

– Pair and Multiple Association Exercises. The application displays elements and these must be related to other ones. The relationship is two elements in the case of pairs and or multiple elements otherwise.
– Categorisation Exercises. In this case, the application displays a number of items and the patient must sort them. The sorting mode depends on the correct order set by the therapist.

3.2 System Architecture

The architecture of the developed system is shown in Fig. 1. On one side, Therapy Design component is handled exclusively by the therapist, as he/she is the one who is allowed to design therapies for each older person. The Therapy Authoring component allows the therapist to create and modify therapies for each patient by using a user-friendly interface. The therapist defines the activities, tasks, and steps needed to develop a therapy [7]. Moreover, the Rule Authoring component allows the control of rehabilitation. The difficulty and accuracy of each exercise, as well as the next step of the rehabilitation process, are determined through controlling the specific variables of the exercise and the vital signs of the user [7]. In addition, there are the Physical/Cognitive Exercise Authoring modules. The Physical Exercise Authoring

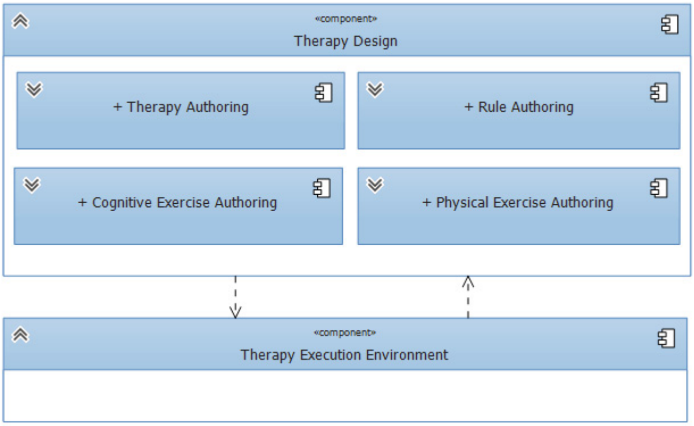


Fig. 1 Component diagram.

module has been presented in previous works [3], and Cognitive Exercise Authoring module will be presented below.

On the other side, this component is responsible for controlling users and system sensors, as well as the execution of rehabilitation exercises. Firstly, achieves the data from all the Kinect sensors and merges them to get an overall picture of the rehabilitation ward [8], [9]. Thus, the system can present the user-data to other system's modules in a unified way. Sencondly, the system monitors the rehabilitation exercises that are carrying out by each user. To achieve a correct control of the therapy execution, the system uses several physiological signals (skin conductance, heart rate, EEG, etc.), captured from external devices, to detect the current state of the patients and act accordingly.

3.3 *Cognitive Rehabilitation Module*

The cognitive rehabilitation module is in charge of supporting therapists and older people in the process of cognitive rehabilitation. Next, the system is shown from the point of view of the roles involved in the process.

Rehabilitation Specialist. The specialist is responsible for designing a specific therapy for a particular patient. Through the Therapy Authoring module he/she is able to design the structure of a therapy, by defining activities, tasks and, finally, steps that are associated with a particular cognitive exercise. Each step has a specific rehabilitation tool that allows the therapist to define a more precise exercise. In this case, the cognitive rehabilitation system is responsible for allowing the therapist to define a specific cognitive exercise.

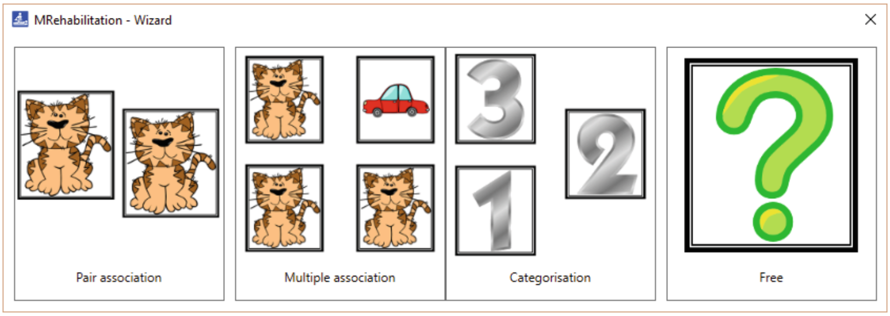


Fig. 2 Type of exercise selection screen.

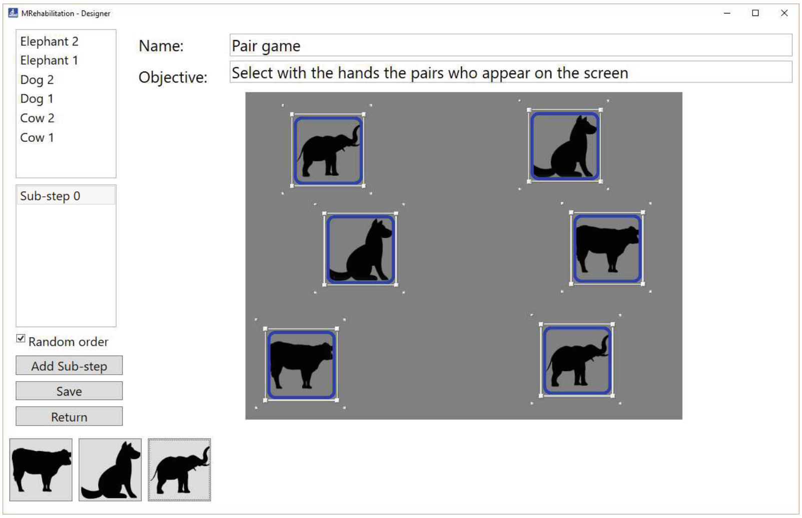


Fig. 3 Exercise edition screen.

When creating an exercise, the four types supported by the system appear (see Fig. 2). After selecting the desired type of exercise, the exercise edition screen is displayed (see Fig. 3). This screen is divided into three areas. At the bottom, the elements necessary to perform the exercise are available to the therapist. On the right side, the name and objective of the exercise, which the specialist must specify, as well as the current work area, are shown. In the current work area, a list of elements that have been added from the bottom area are presented. On the left side, a list of items added to the current workspace are depicted. Besides, there are the 'Sub-steps' that make up the exercise. An exercise may consist of one or more Sub-steps. In each Sub-step, the items displayed can be totally different from the rest of the others.

Each added element is set as a valid, distracting or prohibited item. Valid elements are those on which the user must act when performing an exercise. Activating one

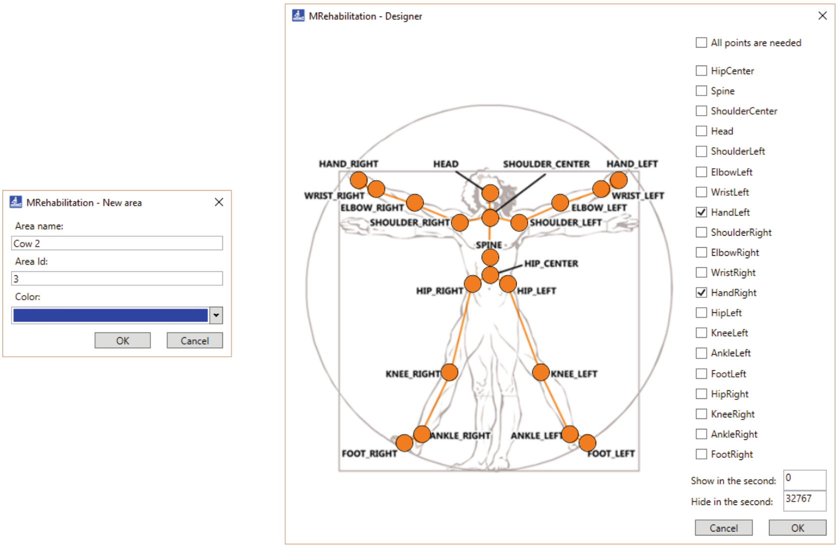


Fig. 4 Zone edition screens.

of them is a hit for the user. Distracting elements are used to add difficulty to the exercise. Each time the user activates one distracting element, it is counted as a cognitive error. Prohibited items represent areas where the user can not enter a part of his/her body. If he/she does so, it is counted as a physical error. The configuration options are shown by double clicking on each area; these are shown in Fig. 4.

All these options enable the rehabilitation specialist to design an exercise as he/she wishes. But they also make a simple exercise to be too time consuming. For this reason, there are some pre-set options for each of the exercises. So, the therapist just needs to change as little as possible.

Older Patients. The patients' part is simpler than the therapist's one. This prevents the user from becoming frustrated and quitting the application, and its use is possible for people with little technical knowledge. The Therapy Execution Environment takes data from previous rehabilitation sessions, the therapy created by the specialist in the Therapy Authoring and the rules of the Rule Authoring modules, and shows the most appropriate rehabilitation exercise to the user.

At the beginning of the exercise, the user can see, at the top of the monitor, the name and the target set by the therapist. These will be read out loud by the system, which will help people with physical or cognitive impairments to get the read text. These should be sufficiently explanatory so that the user does not need any more information for performing the full exercise. In the central part of the screen, the first Sub-step created by the therapist appears, and in its background the figure of the proper person is displayed. This image is captured by a Kinect sensor placed in front of him/her. On the right side of the interface there are also highlighted the number of hits, cognitive and physical failures, the duration of the exercise, and the patient's physiological data (e.g., heart rate or concentration).

4 Discussion and Conclusions

This article has presented a system for computer-assisted cognitive rehabilitation. We have also explained the architecture on which it is based, and which makes it possible for this system to be intelligent and expandable in the future. Because of the wide range of cognitive injuries that may occur, and the large number of exercises that exist to treat them, it is impossible to design a system that covers all of them. Therefore, the constructed system has focused on a small group of exercises (Pair Association, Multiple Association, Categorisation). But it has left the door open to future extensions towards a more complete system.

The development of the tool for the specialist has followed two fundamental principles. On the one hand, the tool is powerful and configurable enough to allow any type of exercise. On the other hand, it is simple enough so that the specialist is not overwhelmed by a very complex tool. This has been achieved by creating a generic tool that allows the user to perform any exercise. A wizard, which provides the most common parameters for each selected exercise, has been added. Thus, the therapist just need to put some elements in the work area. In the case of the ageing adult's application, we have followed the principle of simplicity. Thus, the cognitive load of the application is minimal, since the exercises start automatically. This allows the use of the application to people who do not have all cognitive abilities, or to the ageing adults' group representing the main interest of this application.

The future of the system follows three different lines which are discussed now. The first objective is to conduct a pilot experiment with potential system users and cognitive rehabilitation specialists. Thus, we can get a list of changes and improvements in the system. The second goal is to add new types of exercises that expand the catalogue of cognitive disorders treatments. Thus, the number of potential users should grow. Finally, the third objective is the expansion of the type of devices used for rehabilitation. We propose to use haptic devices for cognitive rehabilitation systems [3] and immersive virtual reality devices [6], among others.

Acknowledgements This work was partially supported by Spanish Ministerio de Economía y Competitividad / FEDER under TIN2012-34003 and TIN2013-47074-C2-1-R grants. Miguel Oliver holds an FPU scholarship (FPU13/03141) from the Spanish Government.

References

1. Chang, Y., Chen, S., Huang, J.: A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. Research in Developmental Disabilities **32**(6), 2566–2570 (2011)

2. Freitas, D., Da Gama, A., Figueiredo, L., Chaves, T., Marques-Oliveira, D., Teichrieb, V., Araújo, C.: Development and evaluation of a Kinect based motor rehabilitation game. Simposio Brasileiro de Jogos e Entretenimento Digital, 144–153 (2012)
3. Oliver, M., Molina, J.P., Montero, F., González, P., Fernández-Caballero, A.: Wireless multi-sensory interaction in an intelligent rehabilitation environment. Ambient Intelligence-Software and Applications **291**, 193–200 (2014)
4. Montero, F., López-Jaquero, V., Navarro, E., Sánchez, E.: Computer-aided relearning activity patterns for people with acquired brain injury. Computers & Education **57**(1), 1149–1159 (2011). doi:10.1016/j.compedu.2010.12.008
5. González, C.S., Toledo, P., Padrón, M., Santos, E., Cairos, M.: TANGO: H: creating active educational games for hospitalized children. In: Management Intelligent Systems, pp. 135–142 (2013)
6. Teruel, M.A., Oliver, M., Montero, F., Navarro, E., González, P.: Multisensory treatment of the hemispatial neglect by means of virtual reality and haptic techniques. Artificial Computation in Biology and Medicine, 469–478 (2015)
7. Rodríguez, A.C., Roda, C., Montero, F., González, P., Navarro, E.: An interactive fuzzy inference system for teletherapy of older people. Cognitive Computation (2015). doi:10.1007/s12559-015-9356-6
8. Oliver, M., Montero-Simarro, F., Fernández-Caballero, A., González, P., Molina, J.P.: RGB-D assistive technologies for acquired brain injury: description and assessment of user experience. Expert Systems **32**(3), 370–380 (2015)
9. Oliver, M., Molina, J.P., Montero, F., González, P., Fernández-Caballero, A.: A multisensor system for positioning of multiple users. In: Proceedings of the XV International Conference on Human Computer Interaction, p. 59 (2014)
10. Gradior (2016). http://www.intras.es/index.php/productos/software-gradior
11. JClic (2016). http://clic.xtec.cat/
12. Bungalow Software (2016). http://www.bungalowsoftware.com/
13. Learning Fundamentals (2016). http://www.learningfundamentals.com/
14. COGRE-HAB (2016). http://www.cogrehab.com/

EEG Mapping for Arousal Level Quantification Using Dynamic Quadratic Entropy

Arturo Martínez-Rodrigo, Beatriz García-Martínez, Raúl Alcaraz, José Manuel Pastor and Antonio Fernández-Caballero

Abstract Developed countries are experiencing a dramatic increase in population ageing. Moreover, it is well known that elderly prefer stay at their homes over other options, increasing this way the medical expenditures. This also involves a major impact on the social and economic balance of the countries. Consequently, an important number of works have been carried out to improve the elderly quality of life and reduce the healthcare costs. However, few efforts have been made in monitoring the mental and emotional states of the ageing adults. This paper introduces a new approach based on the dynamic quadratic entropy for distinguishing different arousal levels. 278 one-minute-length EEG recordings from Dataset for Emotion Analysis using Physiological Signals are used in this study. The results show a decreasing brain complexity under excitement stimuli in central and parietal areas. These findings could be useful to train an emotional system for recognizing some basics emotions.

Keywords EEG mapping · Non-linearity · Elderly · Monitoring · Emotions

1 Introduction

Nowadays, healthcare is the largest area of expenditure in developed countries, mainly due to an increasing population age [1]. Moreover, it is well known that ageing adults prefer staying at their homes over other options, which increases even more medical expenditures [2]. In this regard, it is expected that the percentage of

A. Martínez-Rodrigo(✉) · B. García-Martínez · R. Alcaraz · J.M. Pastor
Instituto de Tecnologías Audiovisuales, Universidad de Castilla-La Mancha, 13071 Cuenca, Spain
e-mail: arturo.rodriguez@uclm.es

A. Fernández-Caballero
Instituto de Investigación en Informática de Albacete, Universidad de Castilla-La Mancha, 02071 Albacete, Spain

© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_23

world population over 65 raises up to 16% in 2050 [3]. This involves a major impact on the social and economic balance of most countries around the world. Bearing all this context in mind, it is not surprising that the World Health Organization considers cost reduction in healthcare systems as one of the most promising challenges in the upcoming years [4].

An important number of works have been carried out in the last years to face this problem, focusing their efforts on the research and development of telemedicine devices or fitness accessories [5], [6], [7]. Novel electronic devices, sensor miniaturization, power saving and efficient communication protocols enable rapid advancements [8]. In this sense, wireless body area networks (WBAN) are emerging as the most reliable and cheapest way to take care of patients who stay at home [9]. WBAN consist of sensors networks which are worn on clothes or implanted on human body, permitting to continuously monitoring some physiological signals. Moreover, WBAN enable physicians to remotely monitor vital signs of patients and provide real-time feedback with medical diagnosis and consultation.

Although many attempts have been made in improving the physical health on elderly, very few efforts have been put in trying to quantify and regulate the elderly mental and emotional state [10], [11]. Ageing adults at home are prone to suffering from loneliness, increasing the probability of falling into mental illness like depression, anxiety or stress [12]. Hence, emotional systems able to interact and recognise human emotions could be the key to face this challenge. However, the inefficiency of machines on recognising human emotional states has been described [13]. Hence, emotion detection models and the development of systems based on emotional intelligence are promising working fields in the near future.

Neurophysiological researchers agree that feeling an emotion leads to changes in the physiological activity, producing variations in the nervous system arousal [14]. Hence, these alterations can be quantified by means of physiological signals which recollect the status of the autonomous central nervous system. Some of the most widely signals used are electromyogram (EMG), electrocardiogram (ECG), electrodermal activity (EDA), skin temperature, among others [15]. However, latest studies are intensifying their efforts in quantifying the brain activity throughout electroencephalography (EEG) recordings, because cognitive processes are primarily generated in our brain [16]. Indeed, it seems more relevant to measure directly over the main source of the emotional processes, rather than using other secondary variables produced indirectly by the brain (muscle contraction, heart rate variability, sweat, etc.). The main drawback of EEG approaches is that the 32-channels standardized 10-20 system requires more computing power to manage all the signals. Moreover, given the non-linearity nature and complex dynamics of EEG recordings, the use of non-linear methodologies [17] is necessary. Indeed, the difficulty to detect any changes in physiological EEG features through traditional lineal methods has been reported [18].

This work aims at studying brain alteration caused during different arousal levels. The excitement or calmness produced by a stimulus can lead to states of relax, stress or anxiety in the elderly. Consequently, the development of effective models able to determine basic feelings is the first step towards an emotion recognition system. A new approach based on dynamic approximated entropy is used in this work to determine the activation of the different areas of the brain in relation with the arousal level.

2 Materials and Methods

2.1 Database

The publicly-available Database for Emotion Analysis using Physiological Signals (DEAP) [13] is used in this work to obtain pre-processed EEG data along with affective ratings given by the subjects under study. The database consists of EEG recordings from thirty two patients (50% both males and females, with mean age equal to 26.9 years). In order to elicit different emotions, the subjects under study visualized forty, one-minute length, music videos with emotional content. Then, participants rated the videos in terms of valence and arousal by using self-assessment manikins. EEG data was acquired at a sampling frequency of 512 Hz, using a standard 10-20 system (32 electrodes) over the scalp. Then, data was processed by a band-pass filtering between 4-45 Hz and down-sampled to 128 Hz. Moreover, eye blink artefacts were removed via independent component analysis. Additional information about the details of this database can be found in [13]. Among the entire database, an EEG recordings subset was formed with those audiovisual fragments where subjects reported an arousal lower than 3 (group I or calmness) and higher than 7 (group II or excitement). Finally, the group under study consisted of 278 samples, where 147 segments belong to group I and 131 segments belong to group II.

2.2 Quadratic Sample Entropy and Its Dynamic Extension

Sample Entropy (*SampEn*) examines a time series for similar epochs and assigns a non-negative number to the sequence, with larger values corresponding to more irregularity in the data [19]. Two input parameters, a run length m and a tolerance window r, must be specified for *SampEn* to be computed. $SampEn(m, r, N)$, being N the length of the time series, is the negative logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, where self-matches are not included in calculating the probability.

Thus, given N data points from a time series $x(n) = \{x(1), x(2), \ldots, x(N)\}$, *SampEn* can be defined as follows:

$$SampEn(m, r, N) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right]. \tag{1}$$

where $B^m(r)$ is the probability that two sequences will match for m points, whereas $A^m(r)$ is the probability that two sequences will match for $m + 1$ points.

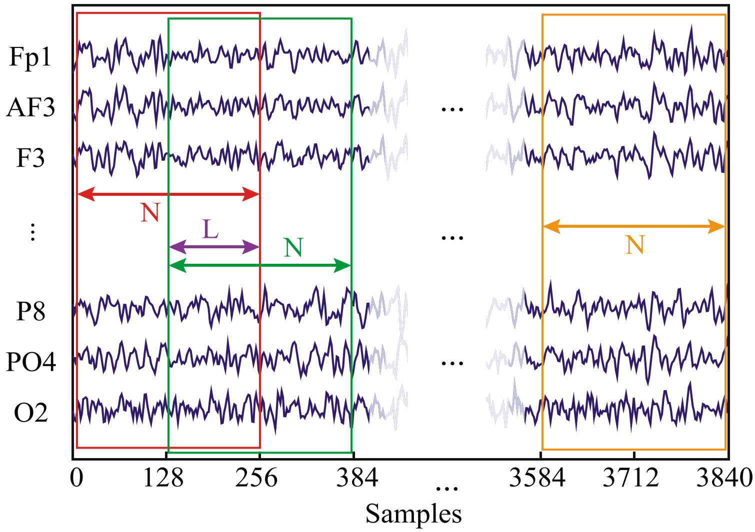


Fig. 1 Dynamic quadratic sample entropy windowing.

Although m and r are critical in determining the outcome of *SampEn*, no guidelines exist for optimising their values [20]. Thus, a slight modification of this entropy metric has been recently proposed to do it more insensitive to the r selection [21]. This new measure, named *Quadratic SampEn* (QSE), allows to vary r as needed to achieve confident estimates of the conditional probability and is defined as:

$$QSE(m, r, N) = SampEn(m, r, N) + \ln(2r). \tag{2}$$

On the other hand, it is worth noting that because the EEG recording is a non-stationary signal [22], a dynamic extension of QSE was used in this work. More precisely, QSE was obtained from a set of consecutive N sample-length windows overlapped L samples, such as Fig. 1 shows. Thus, for a specific EEG recording with M samples in length, dynamic QSE (DQSE) was computed as the average of the QSE values obtained from $\lfloor \frac{M-L}{N-L} \rfloor$ windows, such that:

$$DQSE(m, r, M, N, L) = \text{mean}_{k=1,2,\ldots,\lfloor \frac{M-L}{N-L} \rfloor}\{QSE(m, r, N)_k\}. \tag{3}$$

3 Results

$m = 2$ and $r = 0.25$ times the standard deviation of the original data were used to compute DQSE, because these parameters are the most widely established values for SampEn computation [20]. Additionally, only the last 30 seconds of the EEG

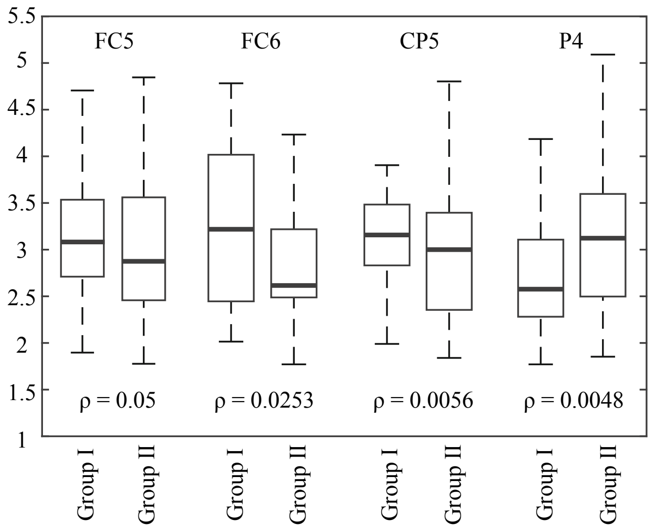


Fig. 2 Boxplot of the DQSE values computed from the most statistically significant recordings, i.e. FC5, FC6, CP5, P4.

recordings (M = 3840 samples) were analysed, similarly to previous works [13]. Finally, after some experiments, $N = 256$ and $L = 128$ samples were chosen, since they reported the best results.

Levene test proved that DQSE distribution was normal and homoscedastic. Consequently, mean and standard deviation values of entropy were computed from EEG channels for the two groups under study. Features reporting statistical differences ($\rho < 0.05$) through a one-way ANOVA test are shown in Fig. 2. As can be appreciated, 4 channels out of 32 reported statistical differences among the two groups. All of them are located in the frontal and parietal zones on the head, being symmetrical to each one. Thus, FC5 and FC6 correspond with left and right channels located in the frontal-central part of the brain, respectively. Similarly, CP5 and P4 correspond with left and right part of central-parietal and parietal zones on the brain, respectively.

As can be observed, average entropy values are lower in FC5, FC6 and CP5 when the arousal level increases. On the contrary, P4 presents higher values of entropy when the arousal level increase. From an statistical point of view, the highest significance was provided by parietal channels CP5 and P4.

In order to provide a global view of brain complexity, EEG maps containing the average entropy for each channels and for all the patients under the two study cases are shown in Fig. 3. In agreement with the results obtained throughout one-way ANOVA test, a low entropy can be observed in frontal-central channels in group B, but not only in FC5, but also in FC1 and FC2. Moreover, central and central-parietal channels experienced also a decreasing in the brain complexity, more concretely C3 and CP5, which are located in the left hemisphere. Indeed, it seems that the left hemisphere experiences, in average, a higher decreasing of brain complexity in

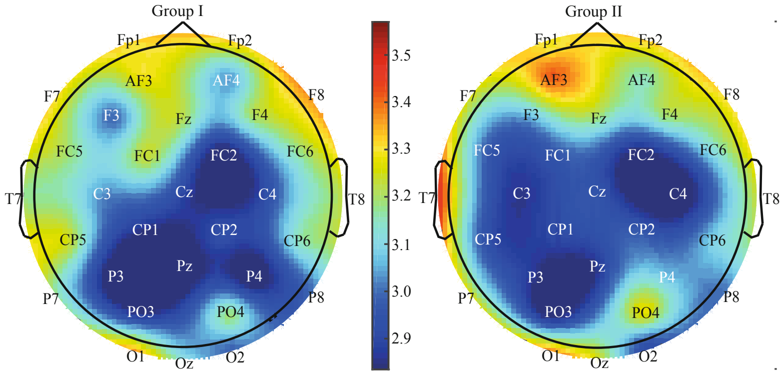


Fig. 3 EEG maps containing the average entropy for each channel and for all the patients under study.

regards to the right hemisphere. On the other hand, parietal channel P4 and temporal channel T7 present a high average brain complexity when the arousal level grows.

4 Discussion and Conclusions

Lately, an important number of studies have put their focus on how the brain works when it is elicited with different kinds of emotions (e.g. [23], [17], [13], [24]). Electroencephalography is emerging as a reliable alternative to the traditional methods based on the quantification of alterations in the autonomous central nervous system. The main reason relays in that brain signals represent actions directly from the source, while other physiological signals are activated by the autonomic nervous system in response to commands originated in the brain. Recent technological advances in sensor miniaturization, together with the development of Software Engineering methodologies, are facilitating new affordable systems for EEG recording and their inclusion into WBAN. The first step towards the development of an emotional system is to find a model to quantify the electro-physiological variations caused by emotional changes. Thus, in this work, a preliminary study is performed. Dynamic quadratic entropy is used to evaluate the non-lineal dynamics presented in the brain when subjects are exposed to different levels of arousal. Although a low-high arousal level can contemplate positive and negative connotations depending on the valence level, our final application is focused exclusively on the excitement or calmness that a stimulus produces in the ageing adult.

Given the statistical significance obtained for some channels, it seems that non-lineal analysis is a good indicator of the brain working dynamics. According to the results, excitement states experience lower entropy levels than calm states. These differences are mainly located at the left hemisphere in central and left parietal channels.

Thus, FC5, FC1, C3 and CP5 show an important average complexity reduction for all the subjects under study (see Fig. 3). Such a reduction represents a decrease in brain system complexity and therefore an increase in regularity. On the contrary, when subjects are under calm events, brain complexity increases, this way elevating the complexity of EEG signals. These findings agree with other works, where a reduction in fractal dimensions in participants under negative emotional stress has been reported [25]. On the other hand, some channels experience an inverse trend, i.e, the brain complexity raises when the arousal level is increased. The existing dissociation between arousal and valence could be the cause of this effect in parietal electrodes [26]. In this regard, it is worth noting that our study is focused exclusively on changes in arousal. Therefore, the variations in valence level could be responsible of the inverse trend in some channels.

In conclusion, the information revealed in this work could be useful to carry out a mathematical model to discriminate some basic emotions such as calm and excitement. The results have shown that central and parietal channels are the most significant channels to focus on, because they experience a global decreasing of the complexity against excitement stimuli. The immediate future work consists in carrying out a discrimination analysis to assess the classification performance of these findings.

Acknowledgements This work was partially supported by Spanish Ministerio de Economía y Competitividad / FEDER under TIN2013-47074-C2-1-R and TIN2015-72931-EXP grants.

References

1. Trustees, M.: Technical review panel on the medicare trustees report, Review of the Assumptions and Methods of the Medicare Trustees
2. Fernández-Caballero, A., Latorre, J.M., Pastor, J.M., Fernández-Sotos, A.: Improvement of the elderly quality of life and care through smart emotion regulation. In: Ambient Assisted Living and Daily Activities, pp. 348–355. Springer (2014)
3. United Nations Department of Economic, World population ageing 2009, vol. 295. United Nations Publications (2010)
4. World Health Organization, et al.: Global health and ageing
5. Pantelopoulos, A., Bourbakis, N.: A survey on wearable systems for monitoring and early diagnosis for the elderly. IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews **1**, 1–12 (2010)
6. Kario, K., Yasui, N., Yokoi, H.: Ambulatory blood pressure monitoring for cardiovascular medicine. IEEE Engineering in Medicine and Biology Magazine **22**(3), 81–88 (2003)
7. Jovanov, E., Milenkovic, A., Otto, C., De Groen, P.C.: A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. Journal of NeuroEngineering and Rehabilitation **2**(1), 6 (2005)
8. Garth, C., Tirthankar, G., Renita, M., Craig, C.: Wireless body area networks for healthcare: A survey. International Journal of Ad hoc Sensor & Ubiquitous Computing **3**(3), 1 (2012)
9. Ullah, S., Kwak, K.S.: An ultra low-power and traffic-adaptive medium access control protocol for wireless body area network. Journal of Medical Systems **36**(3), 1021–1030 (2012)
10. Castillo, J.C., Castro-González, A., Fernández-Caballero, A., Latorre, J.M., Pastor, J.M., Fernández-Sotos, A., Salichs, M.A.: Software architecture for smart emotion recognition and regulation of the ageing adult, Cognitive Computation (in press)

11. Martínez-Rodrigo, A., Zangróniz, R., Pastor, J.M., Fernández-Caballero, A.: Arousal level classification in the ageing adult by measuring electrodermal skin conductivity. In: Ambient Intelligence for Health, pp. 213–223. Springer (2015)
12. Costa, Â., Castillo, J.C., Novais, P., Fernández-Caballero, A., Simoes, R.: Sensor-driven agenda for intelligent home care of the elderly. Expert Systems with Applications **39**(15), 12192–12204 (2012)
13. Koelstra, S., Mühl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I.: Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing **3**(1), 18–31 (2012)
14. Nasoz, F., Lisetti, C.L., Alvarez, K., Finkelstein, N.: Emotion recognition from physiological signals for user modeling of affect. In: Proceedings of the 3rd Workshop on Affective and Attitude User Modelling, Pittsburgh, PA, USA, pp. 1–8 (2003)
15. Valenza, G., Lanata, A., Scilingo, E.P.: The role of nonlinear dynamics in affective valence and arousal recognition. IEEE Transactions on Affective Computing **3**(2), 237–249 (2012)
16. German, W.J.: The hypothalamus and central levels of autonomic function. The Yale Journal of Biology and Medicine **12**(5), 602–603 (1940)
17. Hatamikia, S., Nasrabadi, A.: Recognition of emotional states induced by music videos based on nonlinear feature extraction and som classification. In: 21th Iranian Conference on Biomedical Engineering, pp. 333–337. IEEE (2014)
18. Akar, S.A., Kara, S., Agambayev, S., Bilgic, V.: Nonlinear analysis of eeg in major depression with fractal dimensions. In: 37th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, pp. 7410–7413. IEEE (2015)
19. Richman, J.S., Moorman, J.R.: Physiological time-series analysis using approximate entropy and sample entropy. American Journal of Physiology-Heart and Circulatory Physiology **278**(6), 2039–2049 (2000)
20. Alcaraz, R., Abásolo, D., Hornero, R., Rieta, J.J.: Optimal parameters study for sample entropy-based atrial fibrillation organization analysis. Computer Methods and Programs in Biomedicine **99**(1), 124–132 (2010)
21. Lake, D.E., Moorman, J.R.: Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. American Journal of Physiology-Heart and Circulatory Physiology **300**(1), 319–325 (2011)
22. Abásolo, D., Hornero, R., Espino, P., Poza, J., Sánchez, C.I., de la Rosa, R.: Analysis of regularity in the eeg background activity of alzheimer's disease patients with approximate entropy. Clinical Neurophysiology **116**(8), 1826–1834 (2005)
23. Jirayucharoensak, S., Pan-Ngum, S., Israsena, P.: Eeg-based emotion recognition using deep learning network with principal component based covariate shift adaptation. The Scientific World Journal (2014)
24. Gupta, R., Falk, T.H.: Affective state characterization based on electroencephalography graph-theoretic features. In: 2015 7th International IEEE/EMBS Conference on Neural Engineering, pp. 577–580. IEEE (2015)
25. Hosseini, S.A., Naghibi-Sistani, M.B.: Classification of emotional stress using brain activity. In: Applied Biomedical Engineering, pp. 313–336. INTECH Open (2011)
26. Dolcos, F., Cabeza, R.: Event-related potentials of emotional memory: encoding pleasant, unpleasant, and neutral pictures. Cognitive, Affective, & Behavioral Neuroscience **2**(3), 252–263 (2002)

ARISTARKO: A Software Framework for Physiological Data Acquisition

Arturo Martínez-Rodrigo, José Manuel Pastor, Roberto Zangróniz, César Sánchez-Meléndez and Antonio Fernández-Caballero

Abstract This paper introduces a software framework denominated ARISTARKO. ARISTARKO has been designed and developed to integrate the data acquisition from a wearable physiological data acquisition device with any necessary processing layers in order to classify the acquired signals into a predefined set of emotional states. In this particular article, we use ARISTARKO for the sake of designing an experiment capable of showing a series of images from the well-known IAPS database. The IAPS pictures are labelled with arousal, valence and dominance values. Arousal is classified from the information contained in the database and the physiological signals acquired by the wearable, namely electro-dermal, electrocardiogram and superficial electromyogram activity, and skin temperature.

Keywords Wearable · Physiological data · Software framework · EDA · ECG · EMG · SKT

1 Introduction

Wearable physiological monitoring systems use an array of integrated sensors that continuously acquire and transmit the physiological data to a remote monitoring station. The data captured at the remote monitoring station is correlated to study the overall health status of the wearer. In recent years there has been increasing interest in wearable health monitoring devices. Indeed, a number of experiments have been described on wearable systems for healthcare monitoring (e.g. [1], [2], [3]).

A. Martínez-Rodrigo · J.M. Pastor · R. Zangróniz · C. Sánchez-Meléndez
Instituto de Tecnologías Audiovisuales, Universidad de Castilla-La Mancha, 13071 Cuenca, Spain

A. Fernández-Caballero(✉)
Instituto de Investigación en Informática, Universidad de Castilla-La Mancha, 02071 Albacete, Spain
e-mail: Antonio.Fdez@uclm.es

© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_24

Wearable systems are particularly important to an ever increasingly ageing population [4], [5]. The potential of real-world applications using wearable health monitoring is paramount. Just to name of these, HealthGear, a real-time wearable for monitoring, visualizing and analysing physiological signals, was one of the first systems presented [6]. A more recent paper has reported on the development of a wearable system using wireless biomedical sensors as ubiquitous healthcare service [7].

Nowadays, wearable sensors are being successfully integrated into clothing garments as well as fashion accessories such as hats, wrist bands, socks, shoes, eyeglasses and headphones. These systems often include temperature sensors and accelerometers, which are often used to monitor and classify a person's physical activity. In addition to classifying physical activity, our group is exploring sensors that are fundamental to psychophysiology and understanding of human emotion, in a similar approach to [8].

We feel that it is still a challenge to design and develop wearable data acquisition, processing and transmission systems. These have to be portable and comfortable to wear, provide sustainable battery power and a remote monitoring station [9]. The wearable physiological data acquisition hardware described in this article allows long-term and continuous recording of electro-dermal activity (EDA), electrocardiogram (ECG), superficial electromyogram (EMG) and skin temperature (SKT).

2 Wearable for Physiological Data Acquisition

A first description of the hardware of the wearable for physiological data acquisition has been introduced recently [10]. In Fig. 1, you may observe a picture of the wearable (at the left); some photos of EDA, ECG and EMG devices are shown at the right side of the figure. In the present article, only an excerpt of the previous paper is introduced, as we want to centre here on the software framework for physiological data acquisition.

2.1 Electro-Dermal Activity

The EDA signal morphology corresponds to the superposition of two dominant components. On the one hand, the spontaneous skin conductance (SSC) is the result of an increasing activity in the sympathetic nervous system. It is reflected in the signal as a wave with a variable level of amplitude depending on the intensity and duration of the stimulus [11]. On the other hand, the basal skin conductance (BSC) is related both to the sympathetic nervous system and the dermal characteristics of skin [12]. Considering the different behaviour of these two components, BSC is quantified by calculating the mean over the windowed temporal signal. An impulsive signal is shown when an SSC appears unlike in the basal response. Therefore, the onset and peak boundaries are first detected for each event by using an algorithm capable of

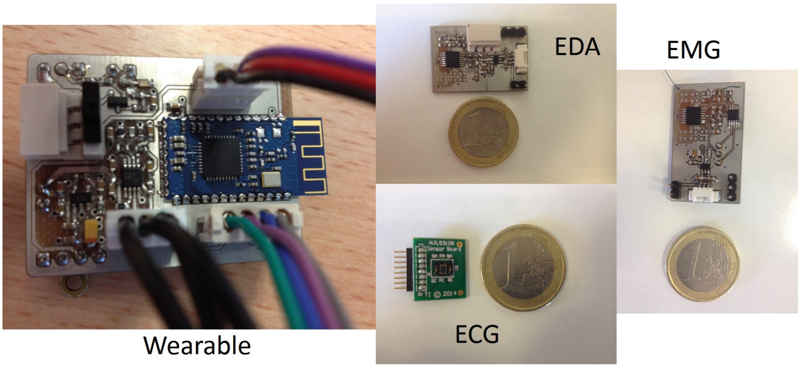


Fig. 1 Wearable for physiological data acquisition.

detecting sudden slope changes [13]. Then, different markers that evaluate the intensity and duration of the events are computed [14]. Thus, duration and intensity are computed by evaluating the temporal distance and magnitude differentiation between onset and peak, respectively.

2.2 Electromyogram

EMG is a measure of the electrical activation of the muscles. It is controlled by the nervous system and produced during muscle contraction. Unfortunately, several kinds of noise signals use to be found when recording an EMG signal, as, for instance, inherent noise in the electrodes, movement artefacts, electromagnetic noise, and cross talk, among others [15]. Consequently, the processing of EMG signals is mandatory to detect muscle activation events, and to reduce or eliminate undesired effects due to noise. A discrete wavelet transform is used, applying a Daubechies's function at decomposition level 4 [16]. Next, the muscle activation event is gotten by applying a simple threshold which determines the existence or absence of EMG activity. Finally, some features involving the muscle activation are calculated. In this work, four temporal markers which involve the EMG integral, the mean absolute value, the root mean square and the variance of the detected EMG events are used [17].

2.3 Electrocardiogram

Electrical activation of the heart takes place by means of the activation of the sinoatrial node which spreads out an electrical impulse from the atrium to ventricles. Heart activity is reflected in the electrocardiogram (ECG) through the P-wave, QRS complex and T-wave, which represent the atrial depolarisation, the ventricular polarization and ventricular depolarisation, respectively. In this regard, while P-wave usually appears in the ECG as a noisy and low amplitude wave, the QRS complex presents the highest amplitude within the ECG, and more specifically the R peak. For this reason, R peak is usually designed to locate the heart cycle within continuous heart activity. In order to track this activity, the R peak is located at each heart cycle and, then, the distance between consecutive R peaks are measured (RR series). This way, the heart rhythm or heart rate is obtained. Nevertheless, given that heart activity is caused by the autonomous nervous system, heart rate variability (HRV), showing the alterations of heart rhythm, is usually computed to evaluate the arousal level of an individual. In our work, the R peaks are firstly located within the ECG by using a second-order derivative algorithm. Then, several temporal features are calculated over the RR-intervals including mean, standard deviation and root mean square.

2.4 Temperature

When human body is under stress, muscles are tensioned, contracting the blood vessels and thus provoking a decrease of temperature [18]. Unlike the aforementioned features, SKT changes are relatively low, such that long-term variability monitoring is necessary. In this work, a mean of temperature, considering a five-minutes temporal window, is used to compute the mean.

3 Framework for Physiological Data Acquisition

ARISTARKO is a software framework developed to integrate the data acquisition from the wearable physiological data acquisition devices with any necessary processing layers in order to classify the acquired signals into a predefined set of emotional states. Fig. 2 shows the main screen of ARISTARKO framework. ARISTARKO offers three options in the menu, namely *File*, *Experiments* and *Database*.

The menu *File* enables the user to load the records of a patient or to add a new patient to the experimentation system. The menu *Experiments* is provided to help the user in designing new experiments. Finally, *Database* menu is thought to connect to a database of previously designed experiments, as well to add a new experiment.

In this article, we have used ARISTARKO for the sake of designing an experiment capable of showing a series of images from the well-known IAPS database [19]. The images contained in IAPS are labelled with arousal, valence and dominance values. Valence and arousal are usually used to obtain faithful information from several emotions, according to the circumplex model of affect [20].

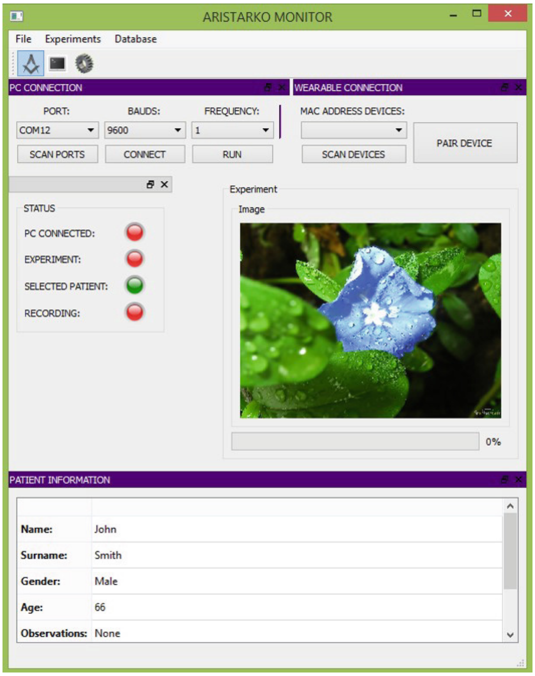


Fig. 2 The ARISTARKO main screen.

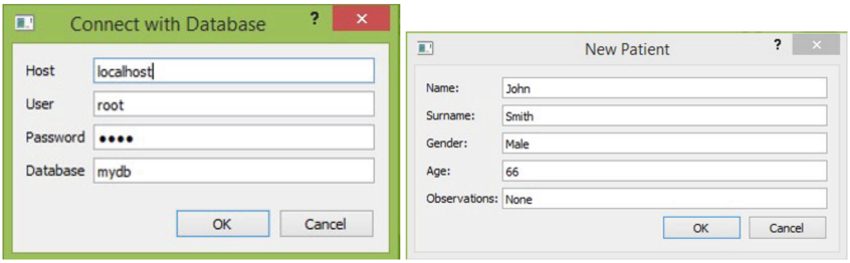


Fig. 3 Left: Connection with database. Right: New patient setup.

3.1 Connecting to a Database

The first step in the design of any experiment is to connect to a database. In this way, it is possible to access to older experiments or to prepare a database to contain a new experiment tat will be designed during a further step. Fig. 3 (left part) shows the layout of the screen for connecting to a database. In order to access the database, the designer has to introduce values for *Host, User, Password* and *Database*.

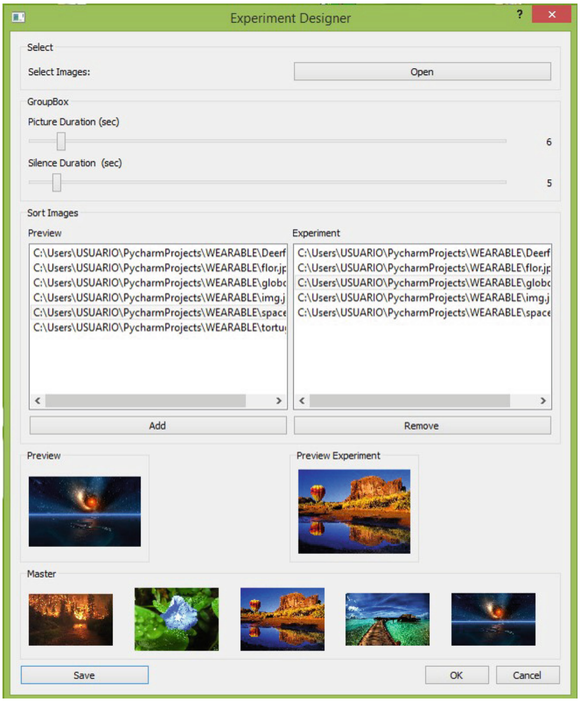


Fig. 4 Experiment designer.

3.2 *Setting up a New Patient*

Should it be necessary to add a new patient (or participant) to the experimental design, the set up is as shown in Fig. 3 (on the right). *Name, Surname, Gender* and *Age* are the minimum parameters to be introduced for a given participant. Obviously, if the patient already exits in the system, he/she is easily accessible.

3.3 *Designing an Experiment*

In this particular case, designing an experiment consists in selecting a series of pictures from the IAPS database that will be used to classify the arousal from the proper database information and the signals acquired from the wearable. Notice in Fig. 4 two very important parameters that are to be tuned for a correct experimentation. These are *Picture Duration* and *Silence Duration*, that is, the time (in seconds) that each image is displayed during the experimentation, and the time that is spent between each pair of images shown to the participant.

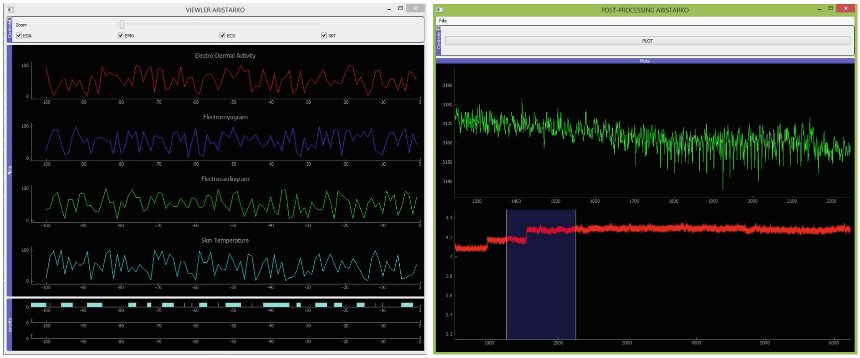


Fig. 5 Left: ARISTARKO viewer. Right: ARISTARKO post-processor.

3.4 Running an Experiment

After designing an experiment, you have to run it. The experiment will be part of the connected database and will be assigned to the current patient. The ARISTARKO viewer (see left side of Fig. 5) enables seeing the continuous drawing of the acquired signals. This way its is possible to verify that the wearable system is working correctly and that there is no anomaly. In order to run the experiment with another patient, it is mandatory to select this participant (obtaining the information from an existing patient or inserting a new one).

3.5 Post-processing the Data

The last step in this implementation of ARISTARKO framework is data post-processing. Fig. 5 (at the right) shows the outcome of the system in terms of the arousal level calculated from the input physiological data of the selected participant. At this moment of the implementation, initial results demonstrate the possibility of classifying the patient's arousal in terms of high, medium and low level, in accordance with the values offered by the IAPS picture database. The classification accuracy is above 85% after experimenting with 30 subjects.

4 Conclusions

This paper has introduced the ARISTARKO software framework. ARISTARKO has been designed to integrate the data acquisition from a wearable physiological data acquisition device with any necessary processing layers in order to classify the acquired signals into a predefined set of emotional states. We have described the main functionalities of the framework.

In this particular approach to ARISTARKO, we have developed a system to design an experiment capable of showing a series of images from the well-known IAPS database. The IAPS pictures are labelled with arousal, valence and dominance values. Arousal is classified from the information contained in the database and the physiological signals acquired by the wearable, namely electro-dermal, electrocardiogram and superficial electromyogram activity, and skin temperature.

Acknowledgements This work was partially supported by Spanish Ministerio de Economía y Competitividad / FEDER under TIN2013-47074-C2-1-R and TIN2015-72931-EXP grants.

References

1. Jung, J., Ha, K., Lee, J., Kim, Y., Kim, D.: Wireless body area network in a ubiquitous healthcare system for physiological signal monitoring and health consulting. International Journal of Signal Processing, Image Processing and Pattern Recognition **1**(1), 47–54 (2008)
2. Hao, Y., Foster, R.: Wireless body sensor networks for health-monitoring. Physiological Measurement **29**(11), 27–56 (2008)
3. Jovanov, E., Milenkovic, A., Otto, C., De Groen, P.C.: A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. Journal of NeuroEngineering and Rehabilitation **2**(1), 6 (2005)
4. Castillo, J.C., Castro-González, Á., Fernández-Caballero, A., Latorre, J.M., Pastor, J.M., Fernández-Sotos, A., Salichs, M.A.: Software architecture for smart emotion recognition and regulation of the ageing adult. Cognitive Computation (2016) (accepted)
5. Fernández-Caballero, A., Latorre, J.M., Pastor, J.M., Fernández-Sotos, A.: Improvement of the elderly quality of life and care through smart emotion regulation. In: Ambient Assisted Living and Daily Activities, pp. 348–355 (2014)
6. Oliver, N., Flores-Mangas, F.: HealthGear: a real-time wearable system for monitoring and analyzing physiological signals. In: Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks, pp. 1–4 (2006)
7. Ogunduyile, O.O., Olugbara, O.O., Lall, M.: Development of wearable systems for ubiquitous healthcare service provisioning. APCBEE Procedia **7**, 163–168 (2013)
8. Fletcher, R.R., Poh, M.Z., Eydgahi, H.: Wearable sensors: opportunities and challenges for low-cost health care. In: 32nd Annual International Conference of the IEEE EMBS, pp. 1763–1766 (2010)
9. Pandian, P.S., Safeer, K.P., Gupta, P., Shakunthala, D.T., Sundersheshu, B.S., Padaki, V.C.: Wireless sensor network for wearable physiological monitoring. Journal of Networks **3**(5), 21–29 (2008)
10. Martínez-Rodrigo, A., Zangróniz, R., Pastor, J.M., Latorre, J.M., Fernández-Caballero, A.: Emotion detection in ageing adults from physiological sensors. In: Ambient Intelligence-Software and Applications, pp. 253–261 (2015)
11. Lidberg, L., Wallin, G.: Sympathetic skin nerve discharges in relation to amplitude of skin resistance responses. Psychopysiology **18**(3), 268–270 (1981)
12. Venables, P.H., Christie, M.J.: Electrodermal activity. Techniques in Psychophysiology, pp. 3–67 (2012)
13. Veltman, J.A., Gaillard, A.W.K.: Physiological indicies of workload in a simulated flight task. Biological Psychology **42**, 323–342 (1996)

14. Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I.: DEAP: a database for emotion analysis using physiological signals. IEEE Transactions on Affective Computing **3**(1), 18–31 (2012)
15. Chowdhury, R., Reaz, M., Mohd, A.M., Bakar, A., Chellappan, K., Chang, T.: Surface electromyography signal processing and classification techniques. Sensors **13**, 12431–12466 (2013)
16. Wei, G., Tian, F., Tang, G., Wang, C.: A wavelet-based method to predict muscle forces from surface electromyography signals in weightlifting. Journal of Bionic Engineering **9**, 48–58 (2012)
17. Hudgins, B., Parker, P., Scott, R.N.: A new strategy for multifunction myoelectric control. IEEE Transactions on Biomededical Engineering **40**, 82–94 (1993)
18. Leijdekkers, P., Gay, V., Frederick, W.: CaptureMyEmotion: a mobile app to improve emotion learning for autistic children using sensors. In: 26th IEEE International Symposium on Computer-Based Medical Systems, pp. 381–384 (2013)
19. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International affective picture system (IAPS): affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL (2008)
20. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology **39**(6), 11–61 (1980)

Vowel Recognition from *RGB-D* Facial Information

José Carlos Castillo, Irene P. Encinar, Alfonso Conti-Morera, Álvaro Castro González and Miguel Ángel Salichs

Abstract One of the main concerns in developed countries is population ageing. Elder people are susceptible of suffering conditions which reduce quality of life such as apraxia of speech, a burden that requires prolongued therapy. Our proposal is intended to be a first step towards automated solutions that assist speech therapy through detecting mouth poses. This work proposes a system for vowel poses recognition from an *RGB-D* camera that provides *2D* and *3D* information. *2D* data is fed into a face recognition approach able to accurately locate and characterize the mouth in the image space. The approach also uses *3D* real world measures obtained after pairing the *2D* detection with the *3D* information. Both information sources are processed by a set of classifiers to ascertain the best option for vowel recognition.

Keywords Apraxia of speech · Visual recognition · Classification · RGB-D

1 Introduction

Population ageing is a recent phenomenon in developed countries which has been drawing attention. People tend to live longer and with high quality of life although in some cases certain conditions such as dementia or Alzheimer prevent the affected ones from living a full and active life. These conditions may involve memory impairment, dyscalculia, agnosia, apraxia and aphasia [12]. In parallel to pharmacological ways of treatment, in recent years social robots have been involved in non-pharmacological therapies to mitigate some of the symptoms. For example, Paro is a socially assistive baby harp seal robot used to support therapeutic interventions for the stimulation of social and communications skills in dementia affected subjects [8].

J.C. Castillo(✉) · I.P. Encinar · A. Conti-Morera · Á.C. González · M.Á. Salichs
Roboticslab, Universidad Carlos III de Madrid, 28911 Madrid, Spain
e-mail: {jocastil,ipencina,aconti,acgonzal,salich}@ing.uc3m.es
http://roboticslab.uc3m.es/

© Springer International Publishing Switzerland 2016
H. Lindgren et al. (eds.), *Ambient Intelligence- Software and Applications (ISAmI 2016)*,
Advances in Intelligent Systems and Computing 476,
DOI: 10.1007/978-3-319-40114-0_25

One limiting cause is *apraxia of speech (AOS)*, a speech motor disorder caused by damage to the brain which hinder saying sounds, syllables, and words. Studies show how this condition can appear even without true language abnormalities. The progressive deterioration of speech production is linked to alterations of mood and dysphagia [3]. Commonly, treatments involve exercises specifically designed to allow the person to repeat sounds over and over and to practise mouth movements for sounds. This exercises are usually supervised by speech therapists, caregivers or even relatives. For this reason, automated solutions that assist speech therapy may help increasing the well-being of people affected by such problems while reducing the burden for caregivers and relatives.

Face recognition systems are able to automatically identify the position of a face and its constitutive elements in images by analysing some specific features [19]. These systems can be also used to isolate and track the different parts composing a face, e.g. the mouth. The bibliography offers a great number of proposals for face detection though not all of them are suitable for an accurate mouth detection since having an adequate set of points composing the mouth is crucial to distinguish different mouth poses. Works such as [10] present good results when detecting faces and locating their main components, although they lack enough detail to allow recognizing different mouth poses. Besides, common problems in face detection are illumination changes and face rotations due to users moving in front of the camera. Another work proposes using local patches around facial landmarks using discriminative areas to reduce illumination problems and in-plane rotation [13]. Other approaches, such as *Stasm* provide acurate mouth recognition, allowing distinguishing the different part of a face [9]. That work uses *Active Shape Models* for locating facial landmark and add a simplified form of *scale-infariant feature transform SIFT* descriptors [7] for template matching. In this sense, using a camera-based system for therapy represent a feasible option. Our goal is to develop a vowel recognition system for *AOS* therapy using information of a commercial-off-the-shelf *RGB-D* camera. In this case, the system will provide not only the mouth position within the face, but also a classification result to identify the user utterance within a set of vowels.

The rest of the paper is organized as follows: Section 2 describes the main steps of the proposal as well as the main classifiers considered in our tests. Section 3 describes the dataset specifically recorded for this work and analyses the results obtained. Finally, the main conclusions are drawn in Section 4.

2 From Face Recognition to Vowel Identification

This work aims to develop a visual vowel recognition system through *2D* and *3D* information. The system described here uses a *Microsoft Kinect*™, which provides *RGB* images and depth data synchronized both in terms of time and field of view. After information acquisition is performed, the system extracts face features in *2D* using the Open Source library *Stasm* [9]. Then, those features are translated into *3D* points which are finally classified to recognize the mouth pose (see Fig. 1).

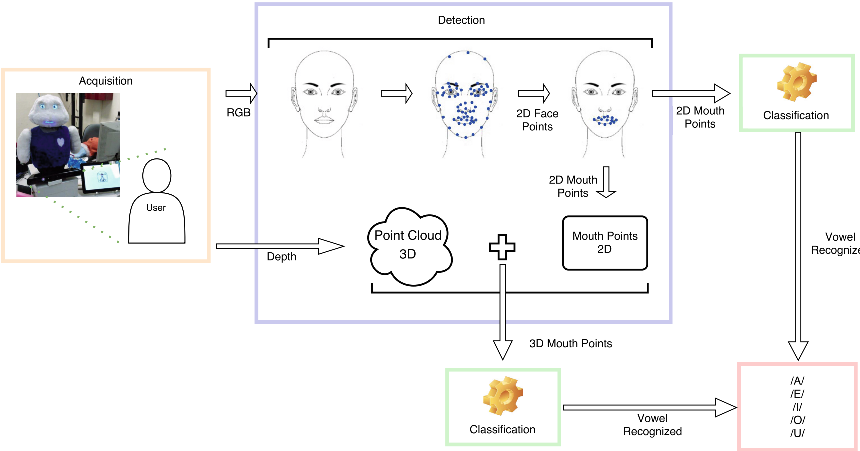


Fig. 1 Main steps of the proposal.

2.1 Process Description

The process is composed by three main steps. The two first ones, namely acquisition and detection, are connected through *ROS* messaging system [14]. The third one, classification, is performed off-line to allow testing several classifiers. The idea of this work is to check the feasibility of different classification techniques towards recognizing vowels. Thus, we have employed *WEKA* [4], a well known data mining tool that allows pre-processing, classification, regression, clustering, association rules, and visualization of data. In our case, several state-of-the-art classifiers are employed to check their suitability in *2D* and *3D* mouth pose recognition.

Data acquisition is performed through a middleware specifically designed to work with *RGB-D* devices, *OpenNI*[1]. Two information flows are generated from the Kinect device: an *RGB* image stream and a point cloud containing depth information. Next, the system processes the *RGB* data to identify the mouth within a detected face using *Stasm*. This library characterizes a face with 77 points of which 18 belong to the mouth. These points are next matched to the depth information from the camera and formatted to be used in the next step, classification.

The last step is to check the feasibility of a set of classifiers in the task of distinguishing different vowel mouth poses from the *2D* and *3D* points composing a mouth in image and real-world coordinates, respectively. The approaches used in this work range from probabilistic techniques to decision trees or neural networks. These classifiers are: *Bayesian networks*, probabilistic models that represent set of variables and their conditional dependences through a directed acyclic graph [2]. *Naive Bayes*, a conditional probabilistic model highly scalable in which the classifier assume that the value of a particular feature is independent of the value of any other feature [17].

[1] OpenNI website: http://openni.ru/

k-nearest neighbors, uses the k closest training examples in the feature space [1]. *Random forest*, an ensemble learning method for classification that operate by constructing a multitude of decision trees [5]. *Multilayer Perceptron (MLP)*, composed by multiple layers of nodes in a directed graph, with each layer fully connected to the next one [16]. *Support Vector Machines (SVM)*, a supervised learning technique that builds a non-probabilistic binary linear classifier [18]. *C4.5* builds decision trees from a set of training data using the concept of information entropy [15]. *1R* is a classifier that learns a one-level decision tree [6].

3 Evaluation of the Proposal

This section describes the dataset we used in this paper to test the different classifiers. To enable comparison, we decided to classify the *2D* mouth points detected by the face recognition algorithm (see Fig. 1) and check those results against the *3D* classification output. The metrics employed are also briefly reviewed in this section as well.

3.1 Dataset Description

Given the specificity of our proposal, recognizing vowels in *RGB-D* information, we were unable to find public datasets that suited our purposes. For this reason, we decided to create our own dataset. The dataset is composed by several video sequences containing different persons pronouncing vowels. More precisely, we collected images from 4 men and 2 women, generating a dataset of 5971 images of which 1089 corresponded to *A*, 1284 to *E*, 1388 to *I*, 1222 to *O*, and 988 to *U*. For our purposes, the dataset just contains images with mouth poses corresponding to the pronunciation of the spanish vowels (/a/ = “ah”, /e/ = “eh”, /i/ = “ee”, /o/ = “oh”, /u/ = “oo”). Intermediate poses are not included since this system will be used for mouth recognition with users holding the same pose steady.

Prior to that, we conducted some tests placing users at different positions in order to test the maximum and minimum distance from the camera at which the detection of the mouth is accurated enough for classification. Figure 2 (a) shows an example of close distance, $0.5m$, in which the mouth detection operates with good resolution (for a face) thus the detection performance is reliable. Figures 2 (b) and (c) show mouth detection at medium distance, $1m$. In this case, the number of pixels composing the mouth clearly decreases, which lead to inaccuracies in the detection. Finally, we tested the feasibility of the detection by placing the user farther from the camera, $2m$. In this case, the detector had to work with a reduced number of pixels and, while the user’s face was still detected, the inaccuracies of the mouth features were too big to consider training a classifier with such points (Fig. 2 (d)). With these data we concluded that our tests should be performed recording users at close or medium distance from the camera to avoid translating detection problems to the classification

(a) (b) (c) (d)

Fig. 2 Detection results at different distances from the camera. 18 dots characterize the mouth. (a) the person is close to the camera and the mouth is correctly detected. (b) the person is at medium distance from the camera and the mouth is correctly detected. (c) the person is at medium distance from the camera and the mouth is not accurately detected. (d) the person is far from the camera and the mouth is not correctly detected.

stage, which might be an issue depending on the final application. In our case this will not be a problem, since vowel detection is thought to operate with users in front of a camera and at a close-range.

3.2 Evaluation Criteria

Some measures widely used by the computer vision community, such as recall, precision and F-score [11], were considered to evaluate the performance of the previously described segmentation algorithms. These measures are calculated as shown in equations (1), (2) and (3), respectively.

$$recall = \frac{TP}{TP + FN} \tag{1}$$

$$precision = \frac{TP}{TP + FP} \tag{2}$$

$$F\text{-}score = \frac{2 \times precision \times recall}{precision + recall} \tag{3}$$

where *TP* (true positives) are the amount of correct mouth detections in the sequence, *FP* (false positives) are the mistaken detections gotten and *FN* (false negatives) are the amount of vowel poses present in the dataset images but not detected.

The precision shows the percentage of true positives with respect to the total number of detections, i.e., the probability of classifications which really correspond to a determined vowel. On the other hand, the recall shows the probability of a specific

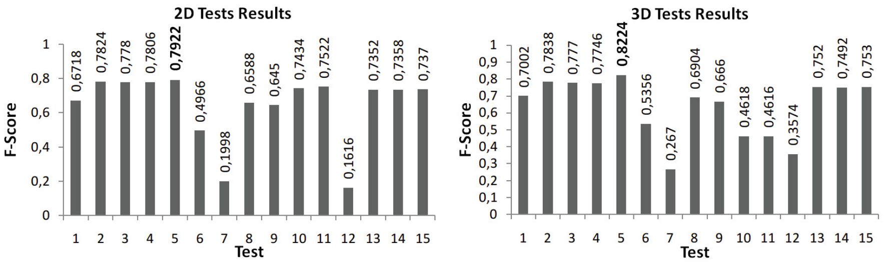


Fig. 3 2D and *3D F-Score* for the classifiers tested. 1) *Bayes Network*; 2) *1-Nearest Neighbors*; 3) *2-Nearest Neighbors*; 4) *4-Nearest Neighbors*; 5) *Random Forest*; 6) *OneR* (Min. Bucket Size = 6); 7) *Naive Bayes*; 8) *MLP* (Learning Rate = 0.2, Training Time = 1000); 9) *MLP*(Learning Rate = 0.3, Training Time = 500; 10) *SVM* (Cache Size = 40, Probability Estimates = False); 11) *SVM* (Cache Size = 40, Probability Estimates = True); 12) *SVM* (Cache Size = 250, Probability Estimates = True); 13) *C4.5* (Confidence Factor = 0,25); 14) *C4.5* (Confidence Factor = 0,1), 15) *C4.5* (Confidence Factor = 0,05)

mouth pose on the image to be really detected. Finally, F-score provides an overall vision of the system performance, considering precision and recall.

3.3 Results

The classifiers described in Section 2.1 are tested with different configurations to ascertain the most suitable for vowel recognition. Some of the classifiers parameters are kept with the default values provided by *WEKA* whereas others have been tuned up as shown in Fig. 3. For all of the tests we used 10-fold cross-validation.

Our initiall assumption was that vowel classification might perform better in *3D* points due to their invariance to distance changes towards the camera since *2D* information extracted from an *RGB* image do not provide information about the absolute distance between the points. Nevertheless, we tested the performance of both options to prove the validity of our hypothesis. Figure 3 left offers the classification F-Score for the *2D* test set. We detected strong variations not only depending on the classifier employed but also on its configuration and the classified vowel. Consequently, results present great variability, having *Random Forest* the best performance for all of the vowels, with an average F-Score of 0, 7922. Other approaches such as *k-nearest neighbors* bring also competitive performance with a 0, 7824 for the best configuration (*1-nearest neighbour*).

Next, the same set of classifiers is run through *3D* mouth points to test the hypothesis that using real world measures is more robust than considering *2D* pixel-level information. As shown in Fig. 3 right F-Score is higher for all classifiers, with *Random Forest* providing again the best classification performance, 0.8224. Again, other approaches, such as *k-nearest neighbors* or *C4.5* provided good performance as well.

4 Conclusions

This paper has presented a comparison of classifiers for vowel recognition from *RGB-D* information. The main steps proposed were acquisition, detection and classification, where Open Source code have been employed to fulfil our goal. *Stasm* and *WEKA* provided the necessary capabilities for detecting and recognizing the mouth pose. The main goal of this work, to introduce a first step towards automated speech therapy using an *RGB-D* commercial sensor, is met considering the results achieved. The *RGB* information provides competitive results, but the classification from *3D* information presents the most accurate performance, with classification values around 80%. Although the system presented in this work does not operate in real-time, we are working on its adaptation to run real experiments with elder people suffering *AOS* in cooperation with therapists, who will ultimately decide whether the obtained results are good enough for our purposes.

Acknowledgements The research leading to these results has received funding from the project called: Desarrollo de robots sociales para ayuda a mayores con deterioro cognitivo (Development of social robots to help seniors with cognitive impairment - ROB- SEN) funded by the Ministerio de Economa y Competitividad (DPI2014-57684-R). This work is partially funded by the MOnarCH European project (FP7-ICT-2011-9- 601033). Also, the research leading to these results has received funding from the RoboCity2030-III-CM project (Robtica aplicada a la mejora de la calidad de vida de los ciudadanos. fase III; S2013/MIT-2748), funded by Programas de Actividades I+D en la Comunidad de Madrid and cofunded by Structural Funds of the EU. The authors would like to thank the Spanish Government for the funds provided by the "Desarrollo de Exoesqueletos Robticos para la Rehabilitacin de Miembro Superior" project (DPI2013-47944-C4-3-R).

References

1. Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician **46**(3), 175–185 (1992)
2. Friedman, N., Geiger, D., Goldszmidt, M.: Bayesian Network Classifiers. Mach. Learn. **29**(2–3), 131–163 (1997)
3. Gerstner, E., Lazar, R., Keller, C., Honig, L., Lazar, G.S., Marshall, R.: A Case of Progressive Apraxia of Speech in Pathologically Verified Alzheimer Disease. Cognitive & Behavioral Neurology **20**(1), 15–20 (2007)
4. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. SIGKDD Explor. Newsl **11**(1), 10–18 (2009)
5. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning, 2nd edn., pp. 587–588. Springer (2008)
6. Holte, R.C.: Very Simple Classification Rules Perform Well on Most Commonly Used Datasets. Mach. Learn. **11**(1), 63–90 (1993)
7. Lowe, D.G.: Object recognition from local scale-invariant features. In: ICCV, vol. 2, pp. 1150–1157 (1999)
8. Marti, P., Bacigalupo, M., Giusti, L., Mennecozzi, C., Shibata, T.: Socially assistive robotics in the treatment of behavioural and psychological symptoms of dementia. In: Biomedical Robotics and Biomechatronics, BioRob, pp. 483–488 (2006)
9. Milborrow, S., Nicolls, F.: Active Shape models with sift descriptors and mars. In: VISAPP, vol. 2, pp. 380–387 (2014)

10. Naruniec, J.: Discrete area filters in accurate detection of faces and facial features. Image and Vision Computing **32**(12), 979–993 (2014)
11. Olson, D.L., Delen, D.: Advanced data mining techniques. Springer Science & Business Media (2008)
12. Ozkan, S., Adapinar, D.O., Elmaci, N.T., Arslantas, D.: Apraxia for differentiating Alzheimers disease from subcortical vascular dementia and mild cognitive impairment. Neuropsychiatric Disease and Treatment **9**, 947–951 (2013)
13. Prabhu, U., Seshadri, K.: Facial Recognition Using Active Shape Models, Local Patches and Support Vector Machines (2009). http://www.contrib.andrew.cmu.edu/~kseshadr/ML_Paper.pdf (Accessed February 6, 2016)
14. Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Ng, A.Y.: ROS: an open-source robot operating system. In: ICRA Workshop on Open Source Software, vol. 3(3.2), p. 5 (2009)
15. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers (1993)
16. Rosenblatt, F.: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, Washington DC (1961)
17. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall (1995)
18. Schölkopf, B., Burges, C.J.C., Smola, A.J.: Advances in Kernel Methods: Support Vector Learning. MIT Press (1999)
19. Li, S.Z., Jain, A.K.: Handbook of Face Recognition. Springer-Verlag (2005)

Author Index