Barbara Guidi
Laura Ricci
Carlos Calafate
Ombretta Gaggi
Johann Marquez-Barja (Eds.)

LNICST

233

Smart Objects and Technologies for Social Good

Third International Conference, GOODTECHS 2017
Pisa, Italy, November 29–30, 2017
Proceedings

EAI

Springer

Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 233

More information about this series at http://www.springer.com/series/8197

Barbara Guidi · Laura Ricci
Carlos Calafate · Ombretta Gaggi
Johann Marquez-Barja (Eds.)

Smart Objects and Technologies for Social Good

Third International Conference, GOODTECHS 2017
Pisa, Italy, November 29–30, 2017
Proceedings

Springer

Editors
Barbara Guidi
University of Pisa
Pisa
Italy

Laura Ricci
University of Pisa
Pisa
Italy

Carlos Calafate
Polytechnic University of Valencia
Valencia
Spain

Ombretta Gaggi
University of Padua
Padua
Italy

Johann Marquez-Barja
University of Antwerp
Antwerp
Belgium

ISSN 1867-8211 ISSN 1867-822X (electronic)
Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering
ISBN 978-3-319-76110-7 ISBN 978-3-319-76111-4 (eBook)
https://doi.org/10.1007/978-3-319-76111-4

Library of Congress Control Number: 2018934333

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

It is a pleasure to welcome everyone to the proceedings of the GOODTECHS 2017 Conference. This was the third edition of the conference and we believe that this conference could become a point of attraction for researchers working in this area. This year the conference was held in Pisa, a historical academic and cultural center.

The main focus of GOODTECHS is the design, implementation, deployment, and evaluation of smart objects and technologies for social good. Clearly, we consider not only these scenarios, but also significant areas where ICT is currently less widespread, hoping that it may represent a societal development opportunity rather than a source for further digital divide.

The conference program comprised technical papers selected through peer reviews by the Technical Program Committee members. After the evaluation process we selected 38 full papers (i.e., a selection rate of 54%). Moreover, we would like to emphasize that GOODTECHS 2017 included two keynote lectures, delivered by experts in their fields: David Plans (University of Surrey, UK) and Stefano Ferretti (University of Bologna, Italy).

We would like to thank the EAI for the support and all the members of the conference committees and the reviewers for their dedicated and passionate work. None of this would happened without the support and curiosity of the authors who sent their papers to this event.

Finally, we would like to encourage current and future authors to continue working in this direction and to participate in events like this conference in order to exchange knowledge and experiences and to make ICT helpful to society.

November 2017

Barbara Guidi
Laura Ricci
Carlos T. Calafate

Organization

GOODTECHS 2017 was organized by the University of Pisa, Italy, in cooperation with ACM SIG Computers and Society, and EAI (European Alliance for Innovation).

Executive Committee

Steering Committee Chair

Imrich Chlamtac	Create-Net Trento, Italy

Steering Committee Members

Ombretta Gaggi	Università degli Studi di Padova, Italy
Pietro Manzoni	Universitat Politècnica de València, Spain
Claudio Palazzi	Università degli Studi di Padova, Italy

General Co-chairs

Barbara Guidi	University of Pisa, Italy
Laura Ricci	University of Pisa, Italy
Carlos T. Calafate	Universitat Politècnica de València, Spain

Technical Program Committee Chair

Ombretta Gaggi	Università degli Studi di Padova, Italy

Publications Chair

Johann M. Marquez-Barja	University of Antwerp, imec, Belgium

Publicity and Social Media Co-chairs

Chaker A. Kerrache	University Amar Telidji of Laghouat, Algeria
Julio Sanguesa	University of Zaragosa, Spain

Web and Local Chair

Andrea De Salve	University of Pisa, Italy

Conference Manager

Katarina Antalova	European Alliance for Innovation, Belgium

Program Committee

Alessandro Lulli	University of Genoa, DIBRIS, Italy
Andrea De Salve	National Research Council, IIT, Italy
Andrea Marino	University of Pisa, Italy
Antonella Molinaro	Università Mediterranea di Reggio Calabria, Italy
Arnoldo Diaz-Ramirez	CETYS Universidad, Mexico
Aniello Castiglione	University of Salerno, Italy
Carlo Mastroianni	Institute for High-Performance Computing and Networking (ICAR-CNR), Italy
Carlos T. Calafate	Universitat Politècnica de València, Spain
Damiano Di Francesco Maesa	University of Pisa, Italy
Emanuele Carlini	National Research Council, ISTI, Italy
Enrique Hernndez Orallo	Universidad Politècnica de València, Spain
Fabio Salice	Politecnico di Milano, Italy
Fabrizio Baiardi	University of Pisa, Italy
Fabrizio Marozzo	University of Calabria, Italy
Francisco J. Martinez	University of Zaragoza, Spain
Hanna Kavalionak	National Research Council, ISTI, Italy
Johann M. Marquez-Barja	University of Antwerp, imec, Belgium
Jorge Herrera Tapia	Universitat Politècnica de València, Spain
Juan Carlos Cano	Universitat Politècnica de València, Spain
Julio Alberto Sangesa Escorihuela	Universidad de Zaragoza, Spain
Michele Tomaiuolo	University of Parma, Italy
Monica Mordonini	University of Parma, Italy
Mohammad Pourhomayoun	California State University, Los Angeles, USA
Paolo Fornacciari	University of Parma, Italy
Rui Campos	INESC Technology and Science, Portugal
Sara Comai	Politecnico di Milano, Italy
Stefano Chessa	University of Pisa, Italy

Sponsoring Institutions

University of Pisa, Italy
ACM SIG Computers and Society
EAI (European Alliance for Innovation)

Contents

Designing a Smart Ring and a Smartphone Application to Help Monitor, Manage and Live Better with the Effects of Raynaud's Phenomenon

Konstantinos Partheniadis and Modestos Stavrakis(✉)

Department of Product and Systems Design Engineering,
University of the Aegean, Syros, Greece
partheniadis.k@gmail.com, modestos@aegean.gr

Abstract. This paper presents the research, the preliminary design stages and an early evaluation of a digital wearable product for monitoring and managing the effects of a chronic disease called Raynaud's phenomenon (RP). The proposed wearable prototype aims at supporting sufferers in their everyday life for managing and preventing RP. The product is composed of three main parts, a physical product of a smart ring, the digital infrastructure of the physical computing subsystem (hardware and software) and an accompanying mobile application.

Keywords: Raynaud's phenomenon · Wearable health technologies
Smart ring

1 Introduction

Digital health is a well-established scientific and industrial domain while the introduction of wearable health technologies that monitor human activity in real-time and provide information for managing health related issues is a relatively new research area that combines, product and industrial design with computing, data analytics and healthcare [1, 2].

The broader scope of digital health and health wearables is to provide patients with technologies (hardware devices, software tools/systems and online services) to better monitor, track and eventually manage their health and wellness related activities. The line between consumer health wearables and specialised medical instruments begins to blur primarily because of the continuous improvement of the former in terms of accuracy, efficiency, aesthetics and form factor. Their use is changing the way people think about their health, but also provide the means to understand and anticipate health progress by (a) providing innovative ways to monitor health and well-being, (b) give greater access to information and (c) support communication and collaboration. According to Lupton [3], digital health technologies and in particular digital health wearables are described as "products that encourage lay people to engage in preventive health activities and improve patient adherence to treatment protocols and their self-management of chronic diseases". Moreover, such technologies afford human

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 1–10, 2018.
https://doi.org/10.1007/978-3-319-76111-4_1

interaction and communication at various levels, including those between healthcare providers and patients but lately, also, among patients and other stakeholders over social network platforms [4].

In this project attention is given to a specific group of people who suffer from RP. In the following paragraphs, we analyse RP and its major characteristics that can be monitored and possibly managed through the use of wearable health technologies. We provide an overview of related projects and present our research, preliminary design stages and an early evaluation of a digital wearable product for monitoring and managing the effects of the chronic disease. Finally, we also outline some insights for future work.

2 Raynaud's Phenomenon

People who suffer from RP very often experience cold fingers, toes and other extremities when they are exposed to cold or stressful situations [5–7]. The typical presentation of a RP event involves the fingers turning white (ischemia), then blue (cyanosis), and finally red (reperfusion). Those events are described as cold 'attacks'. RP appears in 5 to 10% of the world's population with women representing 90% of them. The phenomenon is either described as *Primary* or *Secondary*. *Primary RP* occurs by itself (as a disease) by unidentifiable reasons and can't be cured. *Secondary RP* is considered an expression of another underlying disease which, if identified and managed, it consequently cures RP as well. Both types have common symptoms, but the attacks can differ on symmetry, severity, frequency and duration and can be distinguished with clinical criteria.

Patient education for *prevention* is inseparable part for managing Raynaud's successfully. It is considered that patients should take proper precautions for avoiding direct interactions with cold objects and environments. Multilayered clothing, glove liners, electric gloves, pocket heaters are essential for keeping the core body and extremities warm during the cold months [6]. A healthy eating, non-smoking lifestyle with daily exercising or meditation can help improve circulation and relieve stress in the long-term.

Typical clinical diagnostic investigations include blood count, ESR and ANA analyses, and nailfold capillaroscopy among others. These methods are expensive, available only in specialist centers [8] and require dedicated medical equipment: thermography, arterial Dopplers, large vessel imaging with X-ray, CT or MR angiography. The use of the "cold-challenge test" is often considered the standard measuring/testing mechanism for diagnosing RP while other testing techniques such as the 'distal–dorsal difference' (DDD) in temperature (hypothesis that the tip of a finger is >1 °C colder compared to the dorsum of the hand at room temperature of 30 °C), also help to identify and differentiate RP from other conditions [9].

2.1 Benefits of Monitoring Raynaud's

Monitoring RP as a process has a great potential in managing the disease at various stages in its progress. The benefits can be summarised in the following: *Identify the* *progress of the disease:* Sensors can alert on abnormal body temperature or blood pressure drops that could indicate a Raynaud's attack. Continuous monitoring of severity, duration and frequency of attacks can help in keeping a detailed record of past events automatically. By analysing the stored patterns of occurrence, it is possible to provide insights for the characteristics of future events. Monitoring can also assist in detecting a transition to *secondary* RP as it happens to 2% of patients every year year [10]. Monitoring can also help in identifying specific conditions (e.g. digital ulcerations) that occur during the course of RP and thus provide the grounds for educating users to reduce risks (e.g. scarring or gangrene). *Track and monitor treatment:* Before engaging a patient to a specific treatment plan, the relative risks and benefits must be considered [7]. Thus, monitoring how patients' organisms react to a certain therapy can help identify early benefits or unforeseen risks. Digital health technologies have the potential to monitor the recovery processes that occur with a specific treatment phase and to detect improvements or complications as they arise. Captured data can be analysed and potentially utilised in measuring the effectiveness of a treatment. *Improve self-management:* Adjust behaviour or get motivated for better health e.g. insight to stop smoking. Develop proactive mechanisms to identify and anticipate a 'cold' attack. Engage in a more proactive/prevention lifestyle. Train on thermal biofeedback. Organise post-attack actions. Understand the effects and severity of an attack. *Build a community:* Benefits include the sharing of experiences, confessions, ideas, empirical methods for helping RP. Users can ask questions and exchange information related to their condition (forum). Direct contact with doctors and caregivers. Read news on research, medications, treatments etc. *Develop useful knowledge from raw data:* The psychological nature of RP makes patients evaluate their health status by means of subjective empirical judgments. On the other hand, it is important to gather objectively measured data that can help in assessing the effectiveness of treatment plans. Data acquisition and analysis can provide both users and health professionals with new knowledge on how everyday habits, treatments and behaviors affect RP.

3 Related Work

Today's wearable devices enable real-time monitoring with the aim of self-care and prevention [2, 11]. Heart rate, blood pressure and oxygen saturation are usually measured within the contexts of a physician's office but health wearables are bound to change that [4, 12]. The same measurements can easily and accurately be acquired from today's state-of-the art sensors. Tracking physiological parameters helps provide instant feedback to users, doctors and caregivers by the time a measurement starts getting abnormal. Algorithms can correlate findings and identify patterns that may reveal interesting physiological responses of the body during different activities and environments [4, 13]. Wearable devices can help understand user routines and their baseline norms so as to inform on abnormal changes [13, 14] (e.g. changes in baseline heart rate or skin temperature when waking up). Data analysis can reveal differences among individuals with different health statuses (e.g. those with *Primary* Raynaud's versus those with *Secondary*). They can identify and inform people on habits that have positive or negative impact to their health [14]. Modern wearables can assist users on early medical diagnosis, inform on disease development, track multiple parameters, keep doctors and caregivers updated real-time.

3.1 Related Projects

As of today, there are no low budget, consumer targeted wearable devices (non-medical instruments) or smartphone applications dedicated in tracking continuously and in real-time Raynaud's phenomenon that aim at supporting end-users/patients outside a clinical environment. We identified the following two experimental projects that are related to RP and aim at providing such a service to end-users/patients. In 2002, researchers from John Hopkins University developed a portable device that would wrap around a fingertip to measure RP symptoms at a domestic environment. It featured two temperature sensors for measuring skin and ambient temperatures while the use of a button recorded these values that could later be exported to a computer for further analysis. The device was re-patented in 2005, but no updates on design or user trials are reported ever since [15]. Another monitoring prototype for RP was developed in 2016 as a proof-of-concept for a hackathon event [16]. It featured two temperature sensors; one attached at the wrist and one at the finger in the form of a ring. The project's evaluation mechanism is similar to the 'distal-dorsal difference' hypothesis. The developer's ambition was to create a platform to collect and process everyday data but the project was stalled and no publications were reported.

4 Methodology, Design and Prototyping

4.1 Methodology

The methodology used in this work is based on a mix and match of techniques [17], methods and methodologies used in designing interactive systems and services [18, 19]. It is mainly a multi-methodological approach influenced by interaction design and user experience design that extends to software, systems and service design and industrial design. It encompasses design goals focusing in dealing with product's behavior (physical product and computational, data analytics), visual and physical form, interactivity and user experience. The research and design processes that we followed [20, 21] involved techniques for: project planning; conducting preliminary desktop research (bibliographic, internet); research by collecting data (PACT) [22] and modeling raw information; defining requirements; laying out a basic design framework; designing prototypes and evaluation.

4.2 Design and Prototyping

Design Requirements

According to a recent study, functionality and aesthetics gain surprisingly increased attention among a total of 22 user requirements when it comes to using wearable wellness devices [23]. People prefer lightweight, comfortable, durable and pleasing to use wearable devices that add value to their life. For these reasons, our design requirements for the physical product of the designed ring aim at unobtrusive design and interactions, elegant and minimal aesthetics, durable, waterproof and medical grade materials. Emphasis should be given to the physicality of the interaction where tangible and embodied interactions play a major role and coexist with graphical user interfaces (UIs) that are presented on the mobile application. The device should provide intuitive interaction and user experience and should adapt to the user's everyday activities in a continuous and non-intrusive way. The device must be context-aware and capable of alerting its users when the environment becomes risky, i.e. a 'cold' attack is imminent. Technical requirements include the use of tiny, fast-responding sensors for capturing data (temperature, humidity and skin conductance) and small output actuators for informing and providing feedback (tactile, visual). Data analysis should provide information about user's activity, the context of use, potential psychological distress and the overall RP status by comparing the different patterns of logged data. Other technical requirements include a long-lasting battery and a remote charging capability, as well as means for communicating with a mobile application (e.g. Bluetooth, NFC). The accompanying smartphone application should work as an instant feedback medium to let users monitor and gain insight on their health status. It must keep track of the severity, duration and frequency of RP attacks. The minimum industrial requirements should be related to usability and ergonomics while tangible interactions should focus on delivering maximum integration of the program in everyday life and maximum user experience. At a higher level, the final product (ring and mobile application) must support interactions among patients, doctors and caregivers and provide the means for sharing newly created information to online RP communities. Finally, the product must lead to behaviour change, supporting users for a healthier lifestyle by promoting the values of prevention and active engagement.

Industrial Design

Industrial design is an important stage for this project. It prioritises functionality and aesthetics as an attempt to avoid stigma and misperceptions of assistive technologies. People who use wearable medical devices in their everyday life are often stigmatised and socially discriminated and those who experience RP are not an exemption (cold, pale-colored hands, recovery actions). Therefore, the design of the physical artefact focusses in raising users' self-confidence while minimising the levels of psychological distress caused by social factors (Fig. 1).

The proposed design of the smart ring aims to be indistinguishable from other jewelry. Its form refers to state-of-the-art smart devices with CNC-engineered 0.4 mm titanium shell. The parts are designed for a solid, sealed assembly to allow high levels of water, scratch and drop resistance. The body of the ring is 10 mm wide and 2.4 mm thick, like most metal rings while actual sizes (inside diameter and circumference) may vary. The inner cylindrical part is made of tough, non-slippery medical grade silicon to ensure grasping and avoid skin irritation.

The internal components include three temperature sensors (environment temperature sensor T1, inner 'object-in-contact' temperature sensor T2 and finger temperature sensor T3), a humidity sensor (HM) placed on top and an electrodermal activity (EDA) sensor for monitoring changes in skin conductance. A flex circuit accommodates the microcontroller unit (MCU), a 3-axis accelerometer (ACC), and a

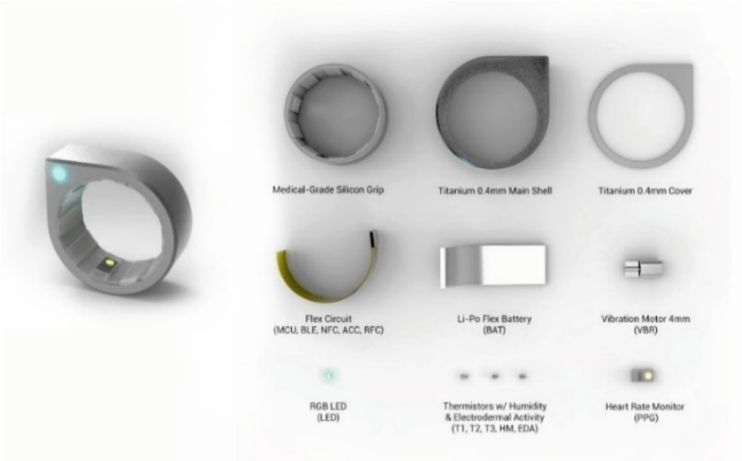


Fig. 1. Physical product of the smart ring and its characteristics, components.

radiofrequency power receiver (RFC) for charging the ring without any cables. The curved Lithium-Polymer (Li-Po) battery is 1.3 mm thick, 26 mm long and 8 mm wide ($1.3 \times 8 \times 26$ mm). Its capacity is estimated to be enough for a two-day battery life (10–20 mAh). The BLE protocol ensures tiny power consumptions and deep sleep modes for when the device is not in use. In the following Table 1 we present how each component contributes to monitoring user's activity and RP symptoms.

Mobile Application and User Experience

The mobile application functions in relation to the ring device and supports automatic sensor data acquisition or manual manipulation of data and other customisations by the end user (personalisation, custom data entry, setting up goals etc.).

It aims in providing useful information for managing and monitoring the course of the disease (log of RP attacks, heartbeat status, temperature, estimated stress levels, user's activity e.g. running, walking, sleeping etc.), it affords online communication mechanisms and synchronisation with third party online applications and services.

The interface is intuitively designed to provide a reliable and pleasing user experience (gradient colors indicate the risk factor, temperature changes can be announced by audible feedback). Device status can be monitored by the user at various levels including proximity and ring's remaining battery life.

Other functionalities include: *Daily insight* for motivating users, *Track an Activity* for detailed analysis of user's activity and goals and in relation to the potential risks for having a RP attack (UI3 at Fig. 3), *Treatment Plan* for managing treatment or medication plans (Info section at Fig. 2 and UI2 at Fig. 3), *Achieve goals* for managing and setting goals related to RP (e.g. proposed by therapist or online by a community, a challenge etc.), *History log* of RP attacks, medicine intakes, tracked activities, achievements and notes are sorted out in a calendar which functions as a detailed history log (UI4, UI5 at Fig. 3), *Share with Doctors* for sharing logged data with doctors, *Make notes* for custom logging and notetaking, *Community of users* for connecting with other users, online communities and services.

Table 1. Electronic components and their role in the smart ring.

	ID	Component type	Role
Input	T1	Environment thermistor	Monitor temp. of environment or object's that covers the finger. (e.g. glove)
	T2	Object-in-contact-thermistor	Monitor the temperature of the things that the user holds or touches with their palm. If hand is on air, it also measures the environment's temperature
	T3	Finger thermistor	Constantly monitors the temperature of the finger. A dangerous temperature drop turns on the PPG sensor to verify the absence of pulse
	EDA	Electrodermal sensor	Skin conductance, sweat rate, psychological distress, water presence
	HM	Humidity sensor	Environment humidity sensor, water presence or absence. Humid environments can trigger a RP attack even when temperature is high
	PPG	Photoplethysmography (PPG) sensor	Measure blood volume changes, monitor heart rate (HR) and heart rate variability (HRV), detect the absence of pulse which indicates a RP attack
	ACC	3-axis accelerometer	Detect user activity (e.g. walk, run) and input gestures (e.g. tap twice for checking RP status or tap three times for checking battery level)
Output	VBR	4 mm vibrator	Receive alerts when environment becomes risky. Feedback to user inputs
	LED	RGB LED	Feedback to user interactions. RGB to represent RP status or battery level
MC unit	MCU	System on Chip (SoC) Microcontroller Unit (MCU)	Microprocessor with integrated Bluetooth Low Energy (BLE) chip and Near Field Communication (NFC) for very low energy consumption
Power	BAT	Flexible (Li-Po) battery	Flexible custom design to fit the product. 2 days of battery life
	RFC	Radio-frequency (RF) wireless charging chip + nest (e.g. Humavox's Thunderlink)	Tiny wireless power receiver located in the ring. Requires a 'nest' where a wireless power transmitter can be implemented to start charging the device when in close proximity
	OTG	On-the-go charger	Accessory to charge ring from smartphone

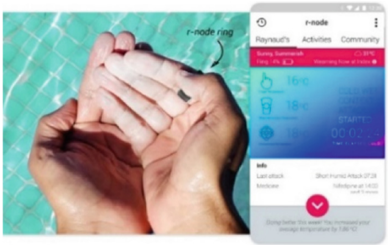


Fig. 2. Wearable ring device and application interface.

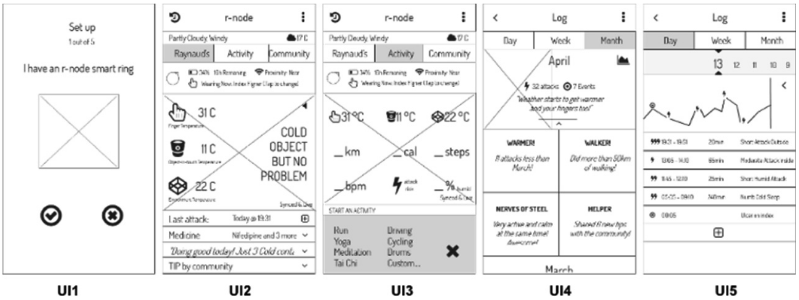


Fig. 3. Low fidelity design of user interfaces. Made using Balsamiq Mockups 3.

The basic interactions and functionalities of the tangible product include a number of gestures for (a) providing feedback in case of an RP attack (upon detection and validation of an attack a system feedback is provided to the user by a specific vibration pattern) (b) a double-tap gesture at the top of the smart ring provides the current status (visual feedback through an RGB led turns a specific color for indicating the current status of the finger) (c) a triple-tap gesture indicates battery level.

To support maximum portability and autonomy the device will charge directly from the user's smartphone using *On-the-go charging* while the *Drop-n-charge* functionality at the 'nest' case will allow wireless charging.

5 Evaluation and Results

At this stage of development three prototypes were designed and evaluated; a *cardboard prototype* for supporting ideation and early ergonomics testing, a *3D printed version of the actual* ring for the evaluation of the final form factor (size, components fitment), ergonomics (sizing, comfort), manufacturing and embodiment and a *working prototype* (wearable ring with electronics and the accompanying Android application) (Fig. 4).

Evaluation was performed in two stages. Initially we did a set of Expert Reviews with double experts for getting fast results, followed by a thorough User Testing. For

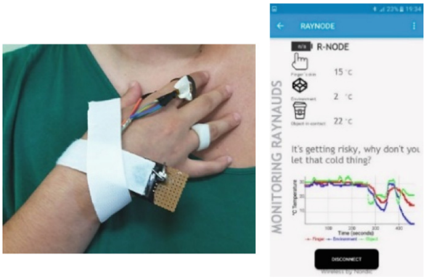


Fig. 4. Working prototype and its application interface.

the later a total of four (n = 4) subjects were recruited to evaluate the prototypes (one diagnosed with RP) under a mixed (CW + HE) cognitive walkthrough and heuristic evaluation usability evaluation method. User subjects followed a specific scenario with a predefined number of tasks (t = 27). Subsequently, a hybrid usability/aesthetics questionnaire was answered.

Evaluation findings show that user subjects believe the application can *help in disseminating scientific knowledge* and *empirical methods* for managing RP, it can *assist communication between patients and health professionals* and *motivate users adopt life-changing behaviours* in the long term. Overall usability and aesthetics received positive comments. We also collected a number of recommendations, both from experts and users, regarding the data analytics and visualisation.

6 Conclusion and Future Work

We identify the potential of a minimal non-obstructive digital health wearable which can monitor and assist in managing the effects of RP at a cost-effective and aesthetically pleasing way for the user. We mainly aim to promote behavior change through minimal everyday interaction with a wearable product and provide accurate information to the end user through data analytics. User experience and interaction design for wearable health technologies, further exploration of the DDD hypothesis and cloud computing for scalable big data analytics are important domains that we intent to further explore in our future designs and research. Instead of designing a single product or application we aim at developing an ecosystem that is composed of several subsystems including a physical product, the online platform and its accompanying cloud and data analytics services [24] and a mobile application for the end users to interact with.

References

1. Pantelopoulos, A., Bourbakis, N.G.: A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans. Syst. Man Cybern. **40**, 1–12 (2010)
2. Chan, M., Estève, D., Fourniols, J.-Y., Escriba, C., Campo, E.: Smart wearable systems: current status and future challenges. Artif. Intell. Med. **56**, 137–156 (2012)

3. Lupton, D.: The digitally engaged patient: self-monitoring and self-care in the digital health era. Soc. Theory Health **11**, 256–270 (2013)
4. Piwek, L., Ellis, D.A., Andrews, S., Joinson, A.: The rise of consumer health wearables: promises and barriers. PLoS Med. **13**, e1001953 (2016)
5. Andrews, J.: Maurice R Raynaud and his protean disease. J. Med. Biogr. **5**, 46–50 (1997)
6. Wigley, F.M.: Raynaud's phenomenon. N. Engl. J. Med. **347**, 1001–1008 (2002)
7. Marcus, S., Weiner, S.R., Suzuki, S.M., Kwan, L.: Raynaud's syndrome. Postgrad. Med. **89**, 171–187 (1991)
8. Herrick, A.L.: The pathogenesis, diagnosis and treatment of Raynaud phenomenon. Nat. Rev. Rheumatol. **8**, 469–479 (2012)
9. Anderson, M.E., Moore, T.L., Lunt, M., Herrick, A.L.: The "distal-dorsal difference": a thermographic parameter by which to differentiate between primary and secondary Raynaud's phenomenon. Rheumatology **46**, 533–538 (2007)
10. Hirschl, M., Hirschl, K., Lenz, M., Katzenschlager, R., Hutter, H.-P., Kundi, M.: Transition from primary Raynaud's phenomenon to secondary Raynaud's phenomenon identified by diagnosis of an associated disease: results of ten years of prospective surveillance. Arthritis Rheum. **54**, 1974–1981 (2006)
11. Chatterjee, S., Price, A.: Healthy living with persuasive technologies: framework, issues, and challenges. J. Am. Med. Inf. Assoc. **16**(2), 171–178 (2009)
12. Li, X., Dunn, J., Salins, D., Zhou, G., Zhou, W., Schüssler-Fiorenza Rose, S.M., Perelman, D., Colbert, E., Runge, R., Rego, S., Sonecha, R., Datta, S., McLaughlin, T., Snyder, M.P.: Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. PLoS Biol. **15**, e2001402 (2017)
13. Chiauzz, E., Rodarte, C., DasMahapatra, P.: Patient-centered activity monitoring in the self-management of chronic health conditions. BMC Med. **13**, 77 (2017)
14. Patel, M.S., Asch, D.A., Volpp, K.G.: Wearable devices as facilitators, not drivers, of health behavior change. Jama **313**, 459 (2015)
15. Wigley, F.M., Wise, R.A., Schwartz, P.D., Lew, A.L., Scott, D.D., Le, B.Q.: Ambulatory surface skin temperature monitor (2005). http://www.google.ch/patents/US6847913
16. Jean, J.: Portable Raynaud Monitoring Device (2016). https://hackaday.io/project/17998-portable-raynaud-monitoring-device
17. Taket, A., White, L.: Partnership and Participation: Decision-Making in the Multiagency Setting. Wiley, Hoboken (2000)
18. Gothelf, J., Seiden, J.: Lean UX: Applying Lean Principles to Improve User Experience. O'Reilly Media, Inc., Sebastopol (2013)
19. Preece, J., Sharp, H., Rogers, Y.: Interaction Design: Beyond Human-Computer Interaction. Wiley, Hoboken (2015)
20. Goodwin, K.: Designing for the Digital Age: How to Create Human-centered Products and Services. Wiley, Hoboken (2009)
21. Cooper, A.: About Face: The Essentials of Interaction Design, 4th edn. Wiley, Indianapolis (2014)
22. Benyon, D.: Designing Interactive Systems: A Comprehensive Guide to HCI and Interaction Design. Pearson Education Canada, Harlow; New York (2010)
23. Rantakari, J., Inget, V., Colley, A., Häkkilä, J.: Charting design preferences on wellness wearables. In: Proceedings of the 7th Augmented Human International Conference 2016 on - AH 2016, pp. 1–4. ACM Press, New York (2016)
24. Zhang, Y., Qiu, M., Tsai, C.-W., Hassan, M.M., Alamri, A.: Health-CPS: healthcare cyber-physical system assisted by cloud and big data. IEEE Syst. J. **11**, 88–95 (2017)

Following the Cuckoo Sound: A Responsive Floor to Train Blind Children to Avoid Veering

Marcella Mandanici[1(✉)], Antonio Rodà[1], Sergio Canazza[1], and Giulia Cavagnoli[2]

[1] Department of Information Engineering, University of Padova, Via Gradenigo, 6, 35131 Padova, Italy
{mandanici,roda,canazza}@dei.unipd.it

[2] Department of Philosophy, Sociology, Education and Applied Psychology, University of Padova, Piazza Capitaniato, 3, 35139 Padova, Italy
giulia.cavagnoli@studenti.unipd.it
http://smc.dei.unipd.it/

Abstract. *Following the Cuckoo Sound* is a responsive floor application designed to train blind children to walk along a straight path through the use of interactive audio. The system, based on computer vision algorithms, is able to estimate the cartesian coordinates of a user as soon as s/he enters the active area. To provide children with an efficient and playful tool which can help them in the task of walking straight, we designed an interactive audio environment based on a cuckoo sound which is altered if the child veers from a central hallway. To obtain preliminary information about the potential of the application, we organized a pilot study involving 6 blind children. Results show an improvement in the stability of the direction in 5 subjects out of 6. Moreover, the great interest among the children for interactive audio suggests that this is a promising investigation field to help wayfinding and orientation.

Keywords: Veering reduction · Gamification · Responsive floor
Interactive audio

1 Introduction

Audio and visual cues provide important information for human walking and wayfinding. But, when the data coming from the visual channel are missing due to visual impairment or total blindness, spatial navigation and wayfinding become a difficult and potentially dangerous task. With the development of computer technologies, many systems have been designed to help blind people to move more independently and without assistance. Basically their structure is formed by a module for the detection of the user's position in space – such as Global Positioning System (GPS), optical motion tracking systems, or others (see e.g. [14,15]) – , a database with environmental data (detailed obstacles map

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 11–20, 2018.
https://doi.org/10.1007/978-3-319-76111-4_2

and data for route planning), and user interface, usually auditory display with spatialized sounds or synthetic speech [1].

1.1 Veering

Despite their efficiency in providing useful data for blind people's navigation, these systems assume the walker's ability to maintain a straight path from one route point to another without changing the orientation in the meanwhile. Which in many cases does not take place, as veering may occur at any time, with severe consequences for instance for blind pedestrians. Veering is a well-known phenomenon that prevents even just blindfolded sighted persons to proceed straight in the absence of some acoustic or physical features such as hallways or sidewalks. Though the reasons of veering are not yet very clear, Kallie et al. [2] postulate that the undetected motor error that produces veering may basically depend on errors in the single step orientation. Particularly, the natural clockwise and counterclockwise rotation of the body during walking would be altered by an excessive lateral placement of the foot that can change the walker's orientation. Assuming that rotation is the primary reason of veering, Guth [3] employed the Anti-Veering Training Device (AVTD) to control the walker's rotation angle. Blind walkers had to follow a simulated crosswalk 2 mt wide and 20 mt long, delimited by ankle-level infrared beams. The system notified the walkers if their rotation angle exceeded the maximum, allowing them to correct their orientation. The results of this training was a reduction of the veering due to the learning of a new pattern of motor output, which persisted for some time after the end of the training.

1.2 An Anti-veering Training Tool for Blind Children

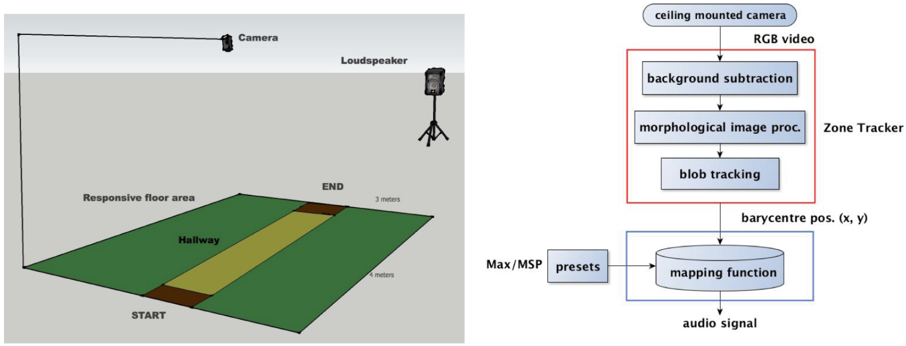
Also if newer systems can take advantage of smartphones built-in sensors to provide blind walkers with useful anti-veering information, exploiting thus the possibilities of an easier and handy technology [4], the existing anti-veering systems are usually intended for adults, and many of them require the user to carry more or less intrusive devices. Moreover, the sounds employed to drive the blind walkers are based only on functional criteria and are not intended to be attractive for the users. In recent years many applications for people with disabilities have been developed considering the potential of a higher involvement of the users through the use of game elements [5]. Motivation, engagement and application effectiveness have proved to produce better results for users in health and wellbeing applications built upon gamification principles with respect to traditional persuasive technology and health games [6]. Gamification can also be successfully employed to foster social interaction for aged people [7], for cognitive rehabilitation, therapy, disease prevention, and assessment [8]. Moreover, it can help in achieving goals due to repetitive and boring exercises, as pinpointed by Rego et al. [9]. Based on these findings, we conceived *Following the Cuckoo Sound*, a responsive floor application based upon a virtual soundscape inspired by a natural environment. The aim of *Following the Cuckoo Sound* is to build a gameful\playful environment devoted not so much to measure the veering of the user but rather to train the children in maintaining a straight route, in a comfortable and funny way. To obtain this result we introduced some game elements in the sound design of *Following the Cuckoo Sound*, with the dual goal of fulfilling for functional purposes and enriching the application's audio output, as described in Sect. 2.1. Beyond this, thanks to the possibilities offered by the *Following the Cuckoo Sound* responsive floor described in Sect. 2.2, the blind child can experience a hands-free and natural interaction, which enhances a sense of freedom and which encourages the environment exploration.

2 Following the Cuckoo Sound

Following the Cuckoo Sound is a responsive floor system designed to train blind children to avoid veering through the use of interactive audio cues. The term *responsive floor* indicates an area where it is possible to track the presence and movement of one or more users, typically employing sensorized tiles or computer vision systems. Our responsive floor is a rectangular surface of 3 × 4 mt placed under the range of a camera hanging from the ceiling of the room. Through the camera the computer can track the movements of a person moving on the floor, allowing thus to link the blind walker's position to the audio output.

2.1 Sound Design and Game Elements

To define the sound characteristics of our environment we relied on the findings of Lewis et al. [10] who have investigated what are the best audio conditions to assist blind walkers. The authors found that a regular intermittent beep was preferred over a continuous signal and that blind walkers preferred to be notified if they were off track, with a different notification if they were off on the right or on the left of the hallway. Moreover they declared to prefer to avoid headphones. Keeping these principles in mind we conceived an ecological soundscape reproducing the sounds of a wood. The choice of a virtual soundscape inspired by a natural environment is the first game element of our application. It provides a sense of immersion which has the function to warm up the blind child's attention and curiosity in navigating the interactive space. In our pilot study we wanted to drive children along a 4 × 0.6 mt hallway put in the middle of the responsive floor, as depicted in Fig. 1a. As soon as the child occupies the *START* area the sounds of the wood are heard together with a cuckoo sound, which is intermittent by nature. The loudspeaker source, put just in front of the hallway, permits a strong and clear guidance on the path. If the child walks inside the hallway the cuckoo sound continues unaltered. If the child gets off track on the right side the cuckoo sound becomes higher in frequency, and if s/he goes on the left side the sound becomes lower. Thus the differences in the sound of the cuckoo are used to mark that s/he is going in the wrong direction and are intended as an invitation to correct it. They could as well be used to assign a score to the child's performance, based on the number of changes reported and on the time



(a) The responsive floor setup (b) The system architecture

Fig. 1. In Fig. 1a is depicted the *Following the Cuckoo Sound* responsive floor setup. The 3 × 4 mt active area (green) has in the middle a hallway (light green) with start and end points (brown). These colors are used only to mark the interactive areas in the present Figure, while in the reality only the green carpet is visible. A hanging camera for motion tracking and a single loudspeaker put in front of the path complete the application's technological setup. In Fig. 1b is depicted the system architecture with its two software modules: the *Zone Tracker* application and the *Max/MSP* patch. (Color figure online)

the child employs to correct her/his route. Finally, as in the video games, when the child reaches the *END* position a final jingle notifies that the target has been successfully reached[1].

2.2 System Architecture

Following the Cuckoo Sound's system is based upon computer vision algorithms. Its technological equipment is composed by a video camera, a long powered USB cable for camera connection, a computer, a green carpet, and a loudspeaker. The carpet's color has been chosen has an optimal background to enhance the camera vision and to avoid excessive light reflection that could altere the visual data and that could be visible even to blind children. The system relies upon two software modules: the *Zone Tracker* application for camera data processing [11,16] and a *Max/MSP* [12] patch for the audio production, as depicted in Fig. 1b. The RGB video data coming from the ceiling mounted camera are processed by the *Zone Tracker* application. After background and morphological image processing, the algorithm defines a blob image, calculates its barycentre and outputs a couple

[1] A video showing *Following the Cuckoo Sound* while being tested at the Robert Hollman Institute (Padova, Italy) can be found at https://youtu.be/yUkPcD1M-OQ. The white tags visible on the carpet in the video were used only to notify the hallway limits and the *START* and *END* positions for clarity of the video viewers. They were removed in the experimental sessions, to avoid that their presence was felt to the touch by the children.

of cartesian coordinates of its position. These data are transmitted via Open Sound Control (OSC) [13] to the *Max/MSP* patch where the mapping functions and audio files are stored.

3 Assessment

To collect some preliminary information about the blind children's behaviour in the *Following the Cuckoo Sound* environment, we organized a pilot study at the Robert Hollman Foundation in Padova (Italy) in October 2016[2]. The study aimed at verifying if, after having walked along the hallway and having followed the cuckoo sound, the children are able to reduce veering in the absence of the audio. Moreover we wanted to check the behavior of the children and to assess their liking to gather suggestions useful to subsequent changes in the system or in the training procedure.

3.1 Subjects

For our pilot study we selected 6 completely blind children of both sexes, aged between 5 and 8 years. As reported in their studies on deviation from a set path, Lewis et al. [10] found a big difference in the response between the test subjects totally blind and partially blind ones. Accordingly, in Table 1 we reported the particular condition of subjects A and E, assuming that this can affect somehow their performance. Subject A was already experienced in the use of the long cane and thus probably had already developed motor cues useful to avoid veering. Subject E was not congenitally blind, which could potentially influence her/his behaviour.

3.2 Procedure

We arranged the responsive floor in the middle of a quiet room approximately of 8×10 mt. Children were introduced in the room one at a time accompanied by their educator. In the room a test assistant and a technical assistant were admitted. Only the test assistant spoke to the children, while all the other people were as silent as possible. Each test was preceded by a short explanatory introduction about the task to accomplish and by an exploratory phase consisting of a walk on the active area to test the reactions of the environment. The organization of each test comprehended the usual three phases of pretest, training and posttest. In the pretest children were accompanied by the test assistant to the starting point of the hallway with their face and shoulders oriented toward the end of the straight path. They had to try to reach the end point for three times without any form of assistance, that is, without any sound produced by the system or voice guidance from the educator. In the second phase the children had to undergo the same procedure but this time they heard the audio output produced by the system which could help them in maintaining the straight path. The posttest was performed exactly in the same way as in the pretest.

[2] http://www.fondazioneroberthollman.it/home-english.html.

3.3 Method

To verify our experimental hypothesis we recorded the data of each subject's walk, obtaining an overall of 9 records for each subject (3 for each phase). As soon as a subject occupies the *START* area the system begins to record the couple of cartesian coordinates of her/his position, tracked as explained in Sect. 2.2, with a frequency of 20 couples per second and stops when the subject reaches the *END* point or when s/he exits the active area. We used these data to plot the path, to measure the veering and to calculate the time employed by the children to travel along the hallway.

Table 1. Overall results of the *Following the Cuckoo Sound* pilot study. For each subject are reported the averages obtained in the three trials of the pre- and posttest with respect to the external area covered by the subject's veering (m^2) and to the time employed (s).

Subject	External area (m^2)			Time (s)		
	Pre	Post	Difference	Pre	Post	Difference
A*	0.00	0.31	−0.31	6.17	6.43	−0.27
B	0.79	0.11	0.69	13.38	7.80	5.58
C	1.19	1.09	0.10	8.93	5.00	3.93
D	1.70	0.74	0.96	9.00	12.80	−3.80
E**	0.14	0.11	0.02	12.33	11.58	0.75
F	0.87	0.45	0.42	12.07	6.27	5.80

*subject alredy trained in the use of the long cane
**subject non congenital blind

3.4 Results

The results of our pilot study are reported in Table 1. As the veering produces a curved area comprised between the straight line of one side of the hallway and the curved line produced by the subject's trajectory, we calculated the external area in m^2 and employed this finding to express the quantity of veering, as can be seen in the example reported in Fig. 2. For each subject we calculated the average of the external area measurements recored in the 3 trials of the pretest and in the 3 trials of the posttest and the difference between the two means. The differences show the effects of the sound walk on the responsive floor. A negative difference means that there is an increase of the external area (corresponding to the veering) in the posttest, while a positive difference shows the amount of the decrease obtained in the posttest trials. We report also the averages of the time (in seconds) employed by the subjects to complete the task in the pretest and in the posttest. Five subjects out of six obtained a decrease of the veering in the posttest with respect to the pretest. Subject A was already well trained in the use of the long cane, which completely prevented her from veering already in the

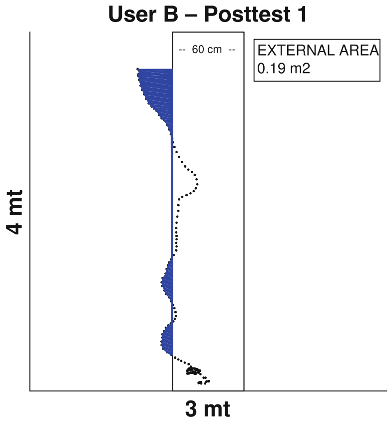
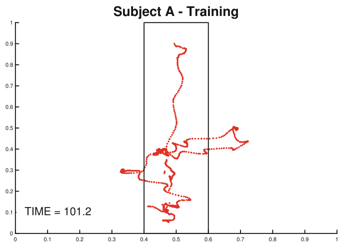
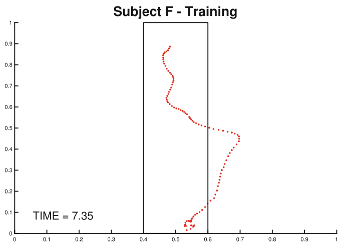


Fig. 2. An example of veering curve plotted employing the subject's coordinates recorded by the system. The rectangular section represents the hallway and the colored area is the surface external to the hallway that expresses the subject's veering (0.19 m^2 in this example). (Color figure online)

3 pretest trials, as reported in Table 1. This led her to consider the training like a game where the veering was done on the purpose of hearing the sound effect, as showed in Fig. 3a. Also the time of this test trial (101.2 s) is much longer if compared to the averages times of the same subject in the pre- and posttest trials (6.17 and 6.43 s respectively). In Fig. 3b is depicted an example of path correction by Subject F as it should happen in consequence of the changes in the audio output. As noticed by Guth [3] and as can be seen in Fig. 4a and b, the training on the responsive floor does not completely eliminate veering, but in the case of subject B noticeably decreases its amount. Subject D obtained the best improvement with respect to the trajectory correction but not with respect to the time employed. In general also a time decrease in the posttest may be

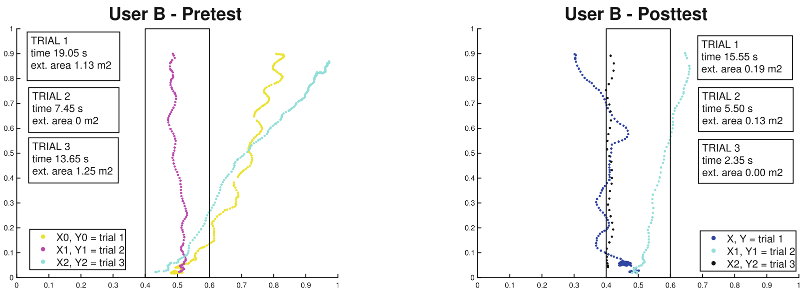


(a) Subject A - training



(b) Subject F - training

Fig. 3. Two training performances from different subjects. In Fig. 3a veering is done on purpose to listen to the sound effect. In Fig. 3b the path correction is made as a consequence of the changes in the audio output. (Color figure online)



(a) Subject B - pretest (b) Subject B - posttest

Fig. 4. The 3 pre- and posttest performances of subject B with noticeable path correction.

interpreted as an improvement of the subject's confidence in walking straight, and thus be considered as a benefit obtained from the training. Subject E, the only non congenitally blind of our sample, performed nearly in the same way as in the pretest, obtaining exactly the same amount of veering and a very small decrease in the time.

3.5 Qualitative Evaluation

Children enjoyed much the experience in the *Following the Cuckoo Sound* environment. Some of them were fascinated by the soundscape produced as soon as they entered the active area and were much more interested in exploring the reaction of the responsive floor than in following the hallway where the cuckoo sound was unaltered (see Fig. 3a). Nine months after the test we found that all the six participants remember the experience and that they are eager to repeat the training. For rating the subjects' liking we used a five items Likert scale survey ranging from 1 (very few) to 5 (very much), obtaining an average score of 4.83.

4 Conclusion

In this paper we presented *Following the Cuckoo Sound*, a training system for the reduction of veering in blind children. Also if the small number of participants prevented us to obtain statistically significant results, our pilot study show that the training with the application seems to reduce the amount of veering and that the observed improvements can be considered as an indication of the potential of this approach. However, some limitations in our experimental setup need to be discussed. The training took place in a single session of 3 trials for an average overall duration of about 1.8 min, a time too short to induce lasting changes. The posttest followed immediately the training and consequently we need to be cautious that the reported decrease in the veering area are due to a

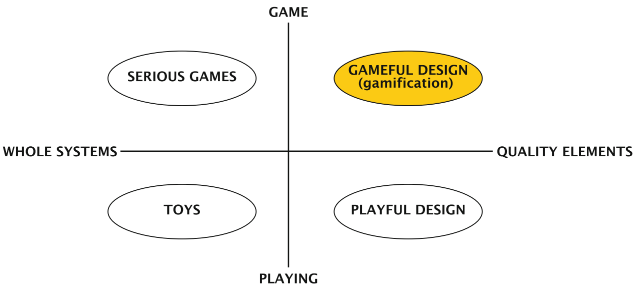


Fig. 5. Conceptual map of gamification applications adapted from Deterding et al. [5]

real improvement in orientation capacity rather than to a memory effect. Thus, further experiments including also a follow-up session are required. Moreover, despite the apparent interest shown by children in exploring the responsive floor area, the short duration of the training did not allow us to verify whether any fatigue or falling interest phenomena occur. Moreover, further research could be done including also longer as well as non-straight paths. Anyway, during the training we had some problem in controlling the behaviour of the children, as the interactive acoustic feedback has a great motivational effect in stimulating the children in the environment discovery. Following the conceptual map depicted in Fig. 5, which cathegorizes games in a bi-dimensional space defined by a *whole-systems* versus *quality-elements* horizontal axis and a *game* versus *playing* (free exploration) vertical axis, we place *Following the Cuckoo Sound* in the *Gameful design* quadrant, that is in the game area. It seems instead that the children interpreted the task of avoiding veering more as a free exploration, placing the application in the *Playful design* area. Perhaps a longer exploratory phase would help in controlling the curiosity of the children and in allowing a tidier training, more similar to a ruled game. Nonetheless this observation leads us to consider the great potential of interactive audio for the design of playful environments which, exploiting the fun and the joy of discovery, can help blind children in achieving a more confident and safe relationship with the space around them.

Acknowledgements. The authors would like to thank *Dr. Vittorina Schoch*, Clinical Director and *Dr. Maria Eleonora Reffo* Administrative Director of the Padua Center of the Robert Hollman Foundation for their cooperation in the organization of the pilot study of the *Following the Cuckoo Sound* application.

References

1. Loomis, J.M., Golledge, R.G., Klatzky, R.L.: Navigation system for the blind: auditory display modes and guidance. Presence: Teleoperators Virtual Environ. **7**(2), 193–203 (1998)
2. Kallie, C.S., Schrater, P.R., Legge, G.E.: Variability in stepping direction explains the veering behavior of blind walkers. J. Exp. Psychol. Hum. Percept. Perform. **33**(1), 183 (2007)

3. Guth, D.: Why does training reduce blind pedestrians veering. In: Rieser, J.J. (ed.) Blindness and Brain Plasticity in Navigation and Object Perception, pp. 353–365 (2007)
4. Panëels, S.A., Varenne, D., Blum, J.R., Cooperstock, J.R.: The walking straight mobile application: helping the visually impaired avoid veering. In: International Conference on Auditory Display. Georgia Institute of Technology (2013)
5. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining gamification. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9–15. ACM (2011)
6. Johnson, D., Deterding, S., Kuhn, K.-A., Staneva, A., Stoyanov, S., Hides, L.: Gamification for health and wellbeing: a systematic review of the literature. Internet Interv. **6**, 89–106 (2016)
7. Gerling, K.M., Masuch, M.: Exploring the potential of gamification among frail elderly persons. In: Proceedings of the CHI 2011 Workshop Gamification: Using Game Design Elements in Non-Game Contexts (2011)
8. McCallum, S.: Gamification and serious games for personalized health. Stud Health Technol Inform **177**(2012), 85–96 (2012)
9. Rego, P., Moreira, P.M., Reis, L.P.: Serious games for rehabilitation: a survey and a classification towards a taxonomy. In: 2010 5th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6. IEEE (2010)
10. Lewis, L., Sharples, S., Chandler, E., Worsfold, J.: Hearing the way: requirements and preferences for technology-supported navigation aids. Appl. Ergon. **48**, 56–69 (2015)
11. Zanolla, S., Canazza, S., Rodà, A., Camurri, A., Volpe, G.: Entertaining listening by means of the stanza logo-motoria: an interactive multimodal environment. Entertain. Comput. **2013**(4), 213–220 (2013)
12. Max/MSP. https://cycling74.com/
13. Wright, M.: Open SoundControl: a new protocol for communicating with sound synthesizers. In: Proceedings of the 1997 International Computer Music Conference, pp. 101–104 (1997)
14. Salvati, D., Canazza, S., Rodà, A.: A sound localization based interface for real-time control of audio processing. In: Proceedings of 14th International Conference on Digital Audio Effects, pp. 177–184 (2011)
15. Salvati, D., Rodà, A., Canazza, S., Foresti, G.L.: A real-time system for multiple acoustic sources localization based on ISP comparison. In: Proceedings of the 13th International Conference on Digital Audio Effects - DAFx-10, pp. 201–208 (2010)
16. Mandanici, M., Rodà, A., Canazza, S.: The "Harmonic Walk": an interactive educational environment to discover musical chords. In: Proceedings of ICMC-SMC, Athens, 14–20 September 2014, pp. 209–216 (2014)

Statistical Features for Objects Localization with Passive RFID in Smart Homes

Kevin Bouchard(✉)

LIARA Laboratory, Université du Québec à Chicoutimi, Chicoutimi, Canada
Kevin.bouchard@uqac.ca

Abstract. Smart homes offer considerable potential to facilitate aging at home and, therefore, to reduce healthcare costs, both in financial and human resources. To implement the smart home dream, an artificial intelligence has to be able to identify, in real-time, the ongoing activity of daily living with a fine-grained granularity. Despite the recent and ongoing improvements, the limitation of the literature on this subject primarily concerns the quality of the information which can be inferred from standard ubiquitous sensors in a smart home. Passive Radio-Frequency Identification is one of the technology that can help improving activity recognition through the tracking of the objects used by the resident in real-time. This paper builds upon the literature on objects tracking to propose a machine learning scheme exploiting statistical features to transform the signal strength into useful qualitative spatial information. The method has an overall accuracy of 95.98%, which is an improvement of 8.26% over previous work.

Keywords: Smart environment · Passive RFID · Indoor localization
Machine learning · Data mining

1 Introduction

World population aging is a situation which most government are now fully aware of the potential consequences, and, therefore, are establishing new policies accordingly to this societal change [1]. While it has been predicted a while ago, the consequence of this continuous aging will be felt for the next few decades to come. One challenge that may be linked directly to this reality is the increasing difficulty to sustain adequate healthcare services to the population. Furthermore, while the population is aging, the life expectancy has also been increasing steadily. Therefore, the active population able to pay for public healthcare is shrinking percentage wise. This problem is quite complex, and researchers, not for profit organizations, and governments now seem to agree that the solutions will be found through innovation from all disciplines involved in the healthcare chain [2].

The miniaturization of technology and the evolution of artificial intelligence (through algorithms improvement and better computer power) enables the researchers to contemplate the implementation of the old smart home dream to alleviate the weight on the fragile healthcare system. Indeed, a large proportion of the direct or indirect healthcare costs can be attributed to the autonomy loss of elders which often result in either a complete care of the person by health professionals (in a long term care establishment or senior housing for non-autonomous) or in a higher frequency of

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 21–30, 2018.
https://doi.org/10.1007/978-3-319-76111-4_3

hospital admission. Many researchers from multiple disciplines, including the members of our laboratory, believe that one of the most important contribution that could be made to relieve the healthcare system would be to enable aging at home through the implementation of smart homes [3]. In this context, a smart home is a technologically enhanced house or residence able to ensure the security of its resident, monitor his health status, and assist him in his activities of daily living (ADLs) in real-time [4]. To do so, smart homes are generally equipped with ubiquitous sensors (passive infrared, electromagnetic contact, etc.), wearable sensors and/or video camera [4–6].

One of the major challenges of implementing a smart home based solution for aging in place is to be able to recognize, in real-time, the ongoing ADLs of the resident [7]. Several methods have been proposed over the past decade, but this endeavor remain problematic due to the low granularity of the current solutions. The granularity, in activity recognition, refers to the level of abstraction provided by the method. For instance, from the lowest to the highest granularity, the same ongoing ADL could be defined as: *Cooking*, *Preparing pasta*, *Preparing shrimp fettucine Alfredo*, or even as the atomic step *Putting fettucine in the boiling water*. While our team at the LIARA laboratory is fairly sensors agnostic [8, 9], we believe that one of the solutions with the highest potential to solve this granularity problem is the passive Radio-Frequency Identification (RFID) technology. The main advantage of passive RFID is that several tags can be installed on daily usage objects in the smart home to enable their tracking in real-time [9]. Therefore, such system could provide highly reliable spatial information to feed an activity recognition algorithm for better granularity.

In this paper, a localization system based on techniques for machine learning/data mining is proposed. The method build upon the work of Bergeron *et al.* [10] which is, in our knowledge, the only example of localization of several objects based on supervised data mining. Indeed, very few authors have worked on the problem of localizing daily usage object, and unfortunately, the best methods for humans/robots tracking often cannot be used straightforwardly [11, 12] because the technology used is too big (require batteries, antennas on the objects, etc.), is too costly, or requires several references points (disposing those in a smart home is not always feasible). As it will be argued further in the paper, daily objects localization is more challenging than human or robot tracking, and the accuracy and precision of the state-of-the-art is still arbitrary. To address this challenge, in this paper, the RFID Received Signal Strength Indication (RSSI) is viewed as a time-series. The research question that was formulated in this project is:"How useful at improving RFID localization methods would be the statistical features commonly used in machine learning?".

The datasets used in this paper were all generated from real data collected in full-scale smart home infrastructures and are available to the scientific community at www.Kevin-Bouchard.com.

2 Related Work

Localization is an old topic of research [13]. Over the years, a plethora of technologies and techniques have been developed and tested for several purposes. This paper could not begin to cover such a vast topic and therefore we encourage the reader to see [14] for a more complete review of the existing works. For wireless technology, there are three widely used techniques. The first one, is the proximity based technique [15]. This technique refer to the association of the tracked object to the closest known point of reference, usually an antenna. The idea is straightforward. The strongest signal among the references determines the position. The reference tags introduced by the LANDMARC system is often categorized as a proximity based technique [16]. The idea is to install tags at known location to use them for accurate tracking of moving tags. The second family are the lateration techniques which use geometric properties to localize an entity. Trilateration is the most often used lateration technique for radio-frequency technologies. The idea is to map the RSSI to a distance measure from the antennas and draw virtual ellipsoid to pinpoint the location at the intersection of few reference points [9]. Finally, the last family of techniques is the learning based methods such as the very popular fingerprinting technique [17]. The fingerprinting technique is usually used in conjunction with a better, more precise, localization system to build a radio map of the environment. The technique is, then, to use the learned map and compare, in real-time, the RSSI to associate the tracked entity to the closest location in a similar fashion than with landmarks. More classical machine learning techniques are less popular [10], since the performance of fingerprinting is usually better. The main drawback is, however, the requirement for the high performance localization system (usually based on ultrasonic sensors) [18].

3 Methodology

In this section, the goal is to explain the methodology that was used to validate the research question formulated in the introduction. While the emphasis of this paper is on the localization of one object in one smart home, the reader should keep in mind the bigger picture, which is about tracking several objects in real-time for ADLs recognition in smart homes for aging at home. Our team has already used the spatial data from passive RFID localization in activity recognition system in the past [19] and improvements in the localization tend translate directly in better activity recognition.

This paper is based on the work of Bergeron *et al.* [10] which was conducted with the LIARA and the DOMUS teams from which the author is a member. The method that was exploited in the aforementioned paper relied, similarly to the literature, on using the raw RSSI signal from the passive tags to perform the localization. In contrast, this project see the RSSI as a time-series. Therefore, despite the low sampling, the localization is performed over a data window, which is an aggregation of many readings. The importance of this work relies on the premise that daily objects localization is more difficult than human/robot localization. The arguments are that daily objects can be very small (e.g.: a spoon, or a fork), numerous (in the kitchen there are several plates, containers, glasses, etc.) and that occlusion will often occur.

3.1 Smart Home

The datasets used in this project were collected in a realistic smart home setting [10]. The smart home is a full scale apartment including a bedroom, a kitchen, a dining room, a living room and a bathroom. It is equipped with 20 polarized directional antennas distributed to cover the entire surface. These antennas are connected to five RFID readers and work on the 928 MHz band as specified by the Canadian Radio-Television and Telecommunications Commision (CRTC). Therefore, they have to be strategically installed to minimize collisions and maximize coverage. Collisions cannot occur among the antennas connected to the same reader since they work on round robin. A derogation can often be obtained through the CRTC to change the band, but since our goal is to use the smart homes for aging in place, this would not be practical. In theory, the RFID system can collect the tags ID up to every 20 ms. However, this is not a real-time system, thus the results are often very different. In practice, it has been observed to be reliably able to collect data under 100 ms (Fig. 1).

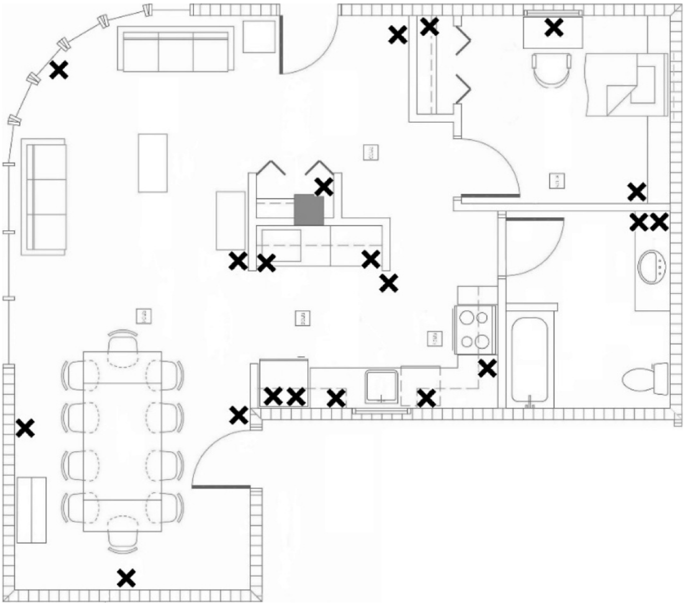


Fig. 1. Aerial view of the smart home and the placement of the antennas from [10].

3.2 Design of Qualitative Zones

In our research context, qualitative information refers to abstracting quantitative information (precise quantities, often continuous variables) into a discrete number of classes or values (e.g.: from GPS coordinates to spatial regions/zones). As I argued in [19], qualitative spatial information is more useful for activity recognition than quantitative. There are two main arguments to this claim. First, it is easier to define (or learn) reasoning rules on qualitative information due to a smaller number of possibilities and better defined classes (weaker interdependencies). Second, qualitative information is an abstraction layer over the quantitative information and therefore it can hide the inherent lack of accuracy. It is especially true in the case of RFID localization.

Precise quantitative methods such as trilateration will always results in different coordinates from iteration to iteration [9], while an abstraction layer of qualitative zones can result into a relatively stable position (albeit being less precise).

To create the qualitative zones, each room is associated with a virtual grid (composed of squares). In the original paper, the qualitative zones were defined heterogeneously. Each room had its own zone's size. The team also made available datasets for alternative zone size in the kitchen and the living room (the datasets are also available at www.Kevin-Bouchard.com). The size of the zones is determined with (1) how much precision is needed in the activity recognition and (2) how many antennas cover the room. They are also limited by the inherent precision of RFID systems. The kitchen and the dining room have respectively 238 and 324 zones of 20 cm by 20 cm. The bathroom and the bedroom have zones of 60 cm by 60 cm and the other rooms have zones of 75 cm by 75 cm.

3.3 Datasets

To understand the properties of the datasets that were used in this project, it is mandatory to first discuss the original datasets published in [10]. The learning was done independently on each of the room. The classes are the set of qualitative zones. There are a total of 673 zones/classes. To collect the data for learning, an object (a plastic bottle of water) equipped with four tags was used. The merging was done through tag selection (the strongest tag was always selected). To understand the impact of using more objects with various shape in real-time, the reader should consult [20]. In the original datasets, fifty readings per classes for each antennas were recorded resulting in 33 650 vectors of twenty RSSI +1 class or 673 000 data. The variation of the RSSI values is bounded between −38 to −69. As a consequence, the datasets have a high number of classes in proportion to the number of possible values.

In this project, the raw RSSI is transformed to obtain one time-series per RFID antenna. Therefore, the transformed datasets are composed of vectors of features computed over the windows of raw RSSI. The features used in the project are discussed in Sect. 3.4. Since the original datasets have a relatively low number of samples per class, the window size is tricky to select. Moreover, for real-time localization of objects, the window must be small enough to avoid lagging in the positioning. Since the system can reliably collect data from 100 ms and more, it seems appropriate to limit the window to between 5 to 10 readings. See the Sect. 4.2 for a better understanding of the impact of selecting a different window size. Finally, in machine learning, there is often the question of sliding windows or not. In our case, the number of samples is actually too low to not use sliding windows for learning. The window slide for each new vector.

3.4 Statistical Features

To exploit the time-series extracted, features were computed over the sliding windows. A wide range of features were added, although the choice was limited by the properties of the datasets. For example, the widely used kurtosis and skewness could not be used because they are extremely affected by sampling [21]. Of course, all statistical features are affected to some extend by the low sampling, but since it could not be predicted how this would translate into the learned models (the sampling could have the effect of making the features highly discriminative), only the statistical features less likely to be impacted were selected. Considering M is the matrix of the data window made of k lines and n columns (the features), the Table 1 describe the features used. There are nine statistical features applied to each time-series and eight applied globally (to all 20 time-series). For instance, the *Mean RSSI* is the sum of all RSSI in a window for an antenna divided by the window size. The *Global Mean RSSI* is the sum of all RSSI in that window divided by the total number of elements in that window ($n * k$). Consequently, the size of each features vector is 189 (20 time-series * 9 statistical features + 8 global features + 1 class = 189).

Table 1. The list of features applied on the time-series.

Mean RSSI	Global mean RSSI	Min RSSI
$\bar{x}_j = \frac{\sum_{i=1}^{k} x_{i,j}}{k}$	$GAvg = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n} x_{i,j}}{n * k}$	$\min(j) = \min_k \{x_{k,j}\}$
Variance of RSSI	Standard deviation of RSSI	Global min RSSI
$Var(X_j) = \frac{1}{k} \sum_{i=1}^{k} (x_{i,j} - \bar{x}_j)^2$	$\sigma_j = \sqrt{Var(X_j)}$	$GMin = \min_n \{\min(j)\}$
Count non-zero	Global mean standard dev.	Max RSSI
$NZ_j = \sum_{i=1}^{k} 1_{\mathbb{R}_{!=0}} (x_{i,j})$	$GStDev = \frac{1}{n} \sum_{j=1}^{n} \sigma_j$	$\max(j) = \max_k \{x_{k,j}\}$
Global NZ	Absolute energy	Global max RSSI
$GNZ = \sum_{j=1}^{n} NZ_j$	$E_j = \sum_{i=1}^{k} x_{i,j}^2$	$GMax = \max_n \{\max(j)\}$
Absolute sum of changes	Global absolute energy	Mean RSSI change
$SC_j = \sum_{i=1}^{k} \lvert x_{i,j} - x_{i-1,j} \rvert$	$GE = \sum_{j=1}^{n} E_j$	$\frac{1}{k} \sum_{i=1}^{k} x_{i,j} - x_{i-1,j}$
Global SC	Global total power	
$GSC = \sum_{j=1}^{n} SC_j$	$Tp = \sum_{i=1}^{k} \sum_{j=1}^{n} x_{i,j}$	

Most of these features are common knowledge, but few of them may need a proper introduction. The *Count Non-Zero*, and by extend the *Global NZ*, count the number of occurrences where the signal was read, or to simply put where the RSSI was different than zero. The *Absolute Energy* is the sum over the squared RSSI values. The *Mean RSSI Change* is the average fluctuation in RSSI that can be expected on the time-series. The *Absolute Sum of Changes* (and *Global SC*) is the sum over the absolute difference between each consecutive RSSI values. Finally, the *Global Total Power* is the sum of all RSSI values over each time-series of the window.

4 Experiments and Results

The first set of experiments that were done had the goal of comparing the results of the features based localization with the original experiments presented in [10]. In the paper, Weka [22], a well-known package of tools and algorithms for data mining, was exploited to learn the models for localization. The default parameters were selected to simplify reproduction of the results and a standard 10-fold cross validation was used to calculate the accuracy. The same experiment was reproduced. The accuracy and performance difference of the algorithm for each of the new datasets (with window size = 5) is compiled in Table 2. The reader should take note that for space purpose only the most interesting algorithms are presented, but in all but one case the accuracy improved over the original method. For K-NN, k = 1 was selected according to the results of testing for k = 1 to 5 in [10]. As the reader can see, in most cases the improvement was significant, especially for algorithm with a lower performance in [10]. Moreover, we can observe that for room with smaller zones and a higher number of classes the improvement was most of the time between 10 to 25%. Overall, the weighted average of the F-measures is 96.097% and the Kappa is 96.062%. The standard deviation of both the F-measures and the Kappa are 5.78% with the lowest score of respectively 0.685 and 0.683.

Table 2. Accuracy for the learning algorithms on each dataset and the performance divergence from [10].

	Dataset	Hall	Living Room	Kitchen	Dining Room	Bedroom	Bathroom	Average
Change in % - Accuracy	CART	99.6%	98.1%	94.0%	95.6%	99.0%	98.4%	**97.4%**
		4.0%	5.7%	20.6%	23.1%	7.5%	8.1%	**11.5%**
	J48 (C4.5)	99.0%	98.8%	96.4%	96.4%	99.1%	98.9%	**98.1%**
		2.2%	6.8%	21.4%	23.1%	6.2%	7.2%	**11.1%**
	1-NN	93.5%	98.2%	68.5%	91.4%	98.8%	95.2%	**90.9%**
		-1.4%	3.4%	-10.3%	18.2%	5.6%	4.8%	**3.4%**
	NaiveBaye	98.9%	99.0%	89.6%	94.9%	99.6%	99.7%	**97.0%**
		2.5%	1.3%	4.6%	20.0%	3.9%	4.6%	**6.2%**
	Random Tree	97.7%	96.0%	89.6%	87.1%	95.8%	95.3%	**93.6%**
		2.1%	6.6%	21.8%	16.0%	6.0%	7.2%	**9.9%**
	Random Forest	99.9%	100.0%	99.8%	98.8%	100.0%	99.8%	**99.7%**
		2.1%	2.4%	10.9%	22.1%	3.5%	3.7%	**7.4%**

4.1 The Impacts of Features

The impact of the features based datasets is more than the accuracy of the learning algorithms. One thing that might be relevant is the complexity of the models built from the new datasets compared to the standard datasets. The tree algorithms such as CART or J48 are useful to do this type of analysis. Random Tree has a random factor that could distort the result, and its complement, Random Forest, is harder to analyze since it is constructed (in our case) of a hundred random trees. When looking at the tree generated by CART and the tree generated by J48 for the *dining room* dataset (324 classes), the size are respectively 1169 and 1187 for the models build with the features. With the raw RSSI of the *dining room*, the tree size is more than the double (CART: 2671 & J48: 2559). This suggest that the new models are more general due to the higher information quality resulting from the features. For simpler room such as the *hall* (X classes), the difference is less staggering. The models size are 47 & 57 for the raw RSSI and 31 & 33 for the features datasets. Notwithstanding its lesser importance, the time to build the models was also compared. Surprisingly, in most cases the time difference between the raw RSSI datasets and the features datasets was not significant. Moreover, it actually improves with Random Tree and (obviously) Random Forest. Finally, the features enable the merging of all the room's datasets together for a unified learning. Although the results are slightly lower, the accuracy is still very high. For example, with J48, the accuracy, Kappa and F-measure are 96.91%. From the 673 different classes, a tree of size 2397 is constructed. With 1-NN, the accuracy drop a little bit more to 82.31%, but for CART it also improves to 95.45% albeit increasing the learning time to 472 s.

4.2 The Impacts of Windowing

The last topic that needs to be discussed is the impact of the window size. Combined datasets were generated for window size 10, 15 and 20 to evaluate how it would affect the accuracy. For J48, the accuracy goes from 96.9% for window size 5 to respectively 99.06%, 99.53% and 99.53%. Obviously, since the accuracy is being already very high, the upside of bigger windows is limited. K-NN is a better candidate to observe this. With window size 5, K-NN accuracy is 82.31%. With the bigger windows, it climb to respectively 93.37%, 97.29% and 98.42%. These observations seem to suggest, as it was expected, that some features become more useful when sampling is more important. However, to use in a real-time localization system, these windows will increase the apparent lag in the positions of the objects. Moreover, with some algorithms such as Random Forest, the accuracy is already very high which hamper the usefulness of increasing the window size. While the learning time of Random Forest is considerable, the resulting model is easily usable in real-time.

5 Conclusion

In this paper, passive RFID was used to localize objects in a smart home for aging in place. The goal was to validate if statistical features could be used to build upon the existing work exploiting machine learning to create automatically localization models.

The experiments demonstrated a gain of 8.26% on average over [10]. In the future, the learned models will be put to test within an activity recognition system at the LIARA smart home in order to confirm that the localization performance gain translate into better recognition of ADLs. In conclusion, we encourage the readers to download the datasets created for this project and use them to pursue their own researches.

Acknowledgments. This project success is the direct consequence of the financial support received from the Université du Québec à Chicoutimi and the National Sciences and Engineering Research Council of Canada.

References

1. Nations, U.: World Population Ageing. Department of Economic and Social Affairs, pp. 1–164 (2015)
2. Moatamed, B., et al.: Low-cost indoor health monitoring system. In: 2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN). IEEE (2016)
3. Al-Shaqi, R., Mourshed, M., Rezgui, Y.: Progress in ambient assisted systems for independent living by the elderly. SpringerPlus **5**(1), 624 (2016)
4. Cook, D.J., et al.: CASAS: a smart home in a box. Computer **46**(7), 62–69 (2013)
5. Hsu, Y.-L., et al.: Design and implementation of a smart home system using multisensor data fusion technology. Sensors **17**(7), 1631 (2017)
6. Bouchard, K., Bouchard, B., Bouzouane, A.: Guideline to efficient smart home design for rapid AI prototyping: a case study. In: International Conference on PErvasive Technologies Related to Assistive Environments, Crete Island, Greece. ACM (2012)
7. Krishnan, N.C., Cook, D.J.: Activity recognition on streaming sensor data. Pervasive Mob. Comput. **10**, 138–154 (2014)
8. Belley, C., et al.: Efficient and inexpensive method for activity recognition within a smart home based on load signatures of appliances. J. Pervasive Mob. Comput. **12**, 1–20 (2013)
9. Fortin-Simard, D., et al.: Exploiting passive RFID technology for activity recognition in smart homes. IEEE Intell. Syst. **30**(4), 7–15 (2015)
10. Bergeron, F., Bouchard, K., Gaboury, S., Giroux, S., Bouchard, B.: Indoor positioning system for smart homes based on decision trees and passive RFID. In: Bailey, J., Khan, L., Washio, T., Dobbie, G., Huang, J.Z., Wang, R. (eds.) PAKDD 2016. LNCS (LNAI), vol. 9652, pp. 42–53. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-31750-2_4
11. Huynh, S.M., et al.: Novel RFID and ontology based home localization system for misplaced objects. IEEE Trans. Consum. Electron. **60**(3), 402–410 (2014)
12. Hekimian-Williams, C., Grant, B., Kumar, P.: Accurate localization of RFID tags using phase difference. In: 2010 IEEE International Conference on RFID IEEE RFID 2010, pp. 89–96 (2010)
13. Dulimart, H.S., Jain, A.K.: Mobile robot localization in indoor environment. Pattern Recogn. **30**(1), 99–111 (1997)
14. Liu, H., et al.: Survey of wireless indoor positioning techniques and systems. IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) **37**(6), 1067–1080 (2007)
15. Song, J., Haas, C.T., Caldas, C.H.: A proximity-based method for locating RFID tagged objects. Adv. Eng. Inform. **21**(4), 367–376 (2007)
16. Ni, L.M., et al.: LANDMARC: indoor location sensing using active RFID. ACM Wirel. Netw. **10**(6), 701–710 (2004)

17. Faragher, R., Harle, R.: Location fingerprinting with bluetooth low energy beacons. IEEE J. Sel. Areas Commun. **33**(11), 2418–2428 (2015)
18. Hightower, J., Borriello, G.: Location systems for ubiquitous computing. Computer **34**(8), 57–66 (2001)
19. Bouchard, K., Bouchard, B., Bouzouane, A.: Spatial recognition of activities for cognitive assistance: realistic scenarios using clinical data from Alzheimer's patients. J. Ambient Intell. Humaniz. Comput. **5**(5), 759–774 (2014)
20. Bouchard, K., et al.: Accurate trilateration for passive RFID localization in smart homes. Int. J. Wirel. Inf. Netw. **21**(1), 32–47 (2014)
21. Joanes, D., Gill, C.: Comparing measures of sample skewness and kurtosis. J. R. Stat. Soc.: Series D (Stat.) **47**(1), 183–189 (1998)
22. Hall, M., et al.: The WEKA data mining software: an update. SIGKDD Explor. Newsl. **11**(1), 10–18 (2009)

Living with Smartwatches and Pedometers: The Intergenerational Gap in Internal and External Contexts

Jayden Khakurel[1(✉)], Susanna Tella[2], Birgit Penzenstadler[3], Helinä Melkas[1], and Jari Porras[1]

[1] Lappeenranta University of Technology, Lappeenranta, Finland
{jayden.khakurel,helina.melkas,jari.porras}@lut.fi
[2] Saimaa University of Applied Sciences, Lappeenranta, Finland
[3] California State University, Long Beach, CA, USA

Abstract. The purpose of this article is to explore and present the range of commonalities and differences between internal and external contexts that influence elderly and younger users' intentions to use commercial off-the-shelf (COTS) smartwatches and pedometers as motivational tools for physical activity. Therefore, this article follows the contextual action theory and the usability evaluation approach, in which "testing" and "inquiry" were applied to 21 younger participants and 13 fit, elderly participants who were in either the pre-contemplation, contemplation, action, or maintenance behavior-change stage. The results revealed no differences in internal context between the target groups due to both the effect and the useful-ness of the external context. However, there were distinctions between the younger and elderly participants regarding external context, especially in certain aspects of device usability, such as font size, touchscreen interaction, interaction technique, and applications installed, which were the core factors that affected the use of COTS smartwatches and pedometers by the study groups. In addition, the external and internal contexts had a cause-effect relationship, which significantly influenced the use of COTS smartwatches and pedometers.

Keywords: Wearable devices · Wearable applications · Smartwatches Pedometers · Elderly · Intergenerational gap · Commercial-off-the-shelf (COTS) Usability

1 Introduction

Much effort has been paid recently to exploring how technologies can promote older adults' well-being and independent living [13]. One area of technology and its user engagement features—such as data, gamification, and content [2]—that has recently become popular among young populations for well-being, and which can be effective to motivate the elderly to be more physically active, is commercial off-the-shelf (COTS) wearable devices. Wearable devices are smart electronic devices available in various forms, worn near or on the body, to sense and analyze physiological and psychological data, such as feelings, movements, heart rate, and blood pressure [12].

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 31–41, 2018.
https://doi.org/10.1007/978-3-319-76111-4_4

This can be done via an application that is either installed on the device or on external devices (e.g. smartphones connected to the cloud) [12]. Wearable devices like activity trackers that measure motion and steps enable users to monitor their behavior and could support a healthier lifestyle [19]. They feature different degrees of usability and a varying range of user experiences [12]; the International Organization for Standardization [9] defines "usability" as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context. Currently, this definition doesn't apply to the elderly, as they have a more difficult relationship with COTS devices than their younger counterparts [3], primarily because hardware and software have not been designed to suit their physical or mental abilities [14], which can discourage the elderly's adoption of devices such as smartwatches and pedometers as tools to perform physical activities.

Despite growth in the use of COTS smartwatches and pedometers, few studies have drawn technological comparisons between the elderly and their younger counterparts [5, 21, 30]. However, no studies have considered how elderly and younger users' perceptions of, and usability challenges associated with, COTS smartwatches and pedometers varies and affects their adoption due to contexts. Context encompasses an internal and external context [7, 20]. The internal context describes users' state and consists of internal parameters of human experience and activity [7] such as emotional responses (e.g. a decrease in user satisfaction and motivation [20]) and manifested behavioral responses such as an increase in errors, in reactions, or in inefficient or inappropriate activities [20]. The external context describes the environmental state and consists of proximity to objects [7], such as devices and their associated applications. To fill the research gap, the present study explores the divide in contexts (internal and external) that appears between target user groups (fit elderly users and younger users) while using the same COTS smartwatches and pedometers and participating in the same usability experiments. Thus, the research question (RQ) is: "Which internal and external contexts can obstruct the use of COTS smartwatches and pedometers among both elderly and younger users while using them as motivational tools for physical activities?" To answer this RQ, we will follow the contextual action theory (CAT) presented by Stanton et al. [20] and a usability evaluation method [10] to explain human action in terms of coping with technology within a context. The outcomes of this study identify challenges associated with wearables that need to be addressed by stakeholders, including device manufacturers, researchers, and caregivers, to enhance user experience, by understanding factors relating to internal and external context.

2 Related Work

Gregor et al. [6] classified the elderly into two categories: fit, who do not appear—nor would consider themselves—disabled, but whose functionality, needs, and wants are different to those they had when they were younger; and frail, considered to have one or more "disabilities," often severe, and who will have a general reduction in many functionalities and require general assistance from caregivers or relatives. Chodzko-Zajko et al. [4] concluded that regular exercise by the fit elderly can have significant psychological and cognitive benefits for their health, which is consistent with the 2008 Physical Activity Guidelines for Americans [25]. Nelson et al. [15] and Tudor-Locke et al. [24] pointed out that regular physical activity can help both the fit and frail elderly in preventing and treating disease and reducing the risk of developing other chronic diseases, premature mortality, functional limitations, and disabilities.

The elderly population is the least physically active of any age group [25], and little is known about how they can be motivated to engage in physical activities to enhance their well-being and independent living. Siek et al. [21] found no major differences in performance between older and younger users when physically interacting using mobile computing devices and completing tasks that are not complex and don't require maximum cognitive effort. However, they found differences in terms of preferences, such as for font sizes. Fukuda and Bubb [5] compared younger and elderly users' web use and found differences related to navigational behavior due to the decline of elderly users' visual and fine motor functions. Meanwhile, Zhou et al. [30] concluded that ageing has significant negative effects on performance and accuracy.

3 Study

Methodological approach. To enhance our understanding of commonalities and differences among elderly and younger participants using the same device in the same experiments, CAT and a usability evaluation method [10] form the foundation of this methodology. According to CAT, human behavior can be segmented into actions by assuming, attributing, or reporting a goal for the behavior [29]. Stanton et al. [20] pointed out that CAT explains human actions in terms of coping with technology within a context, with five phases associated with contextual actions: (i) actual demands and resources are presented to the user, which comprise the design of the device, the tasks to be performed on the device, environmental constraints (e.g. time) and so on; (ii) appraisal of those demands and resources by the actor; (iii) a comparison of the perceived demands and resources; (iv) possible degradation of pathways; and (v) the effects of these responses on the interaction with the devices.

The type of internal and external data we gather from action is also dependent on the data-gathering procedures [29]. Therefore, we applied a usability evaluation method composed of a series of well-defined activities to collect data related to the interaction between the end user and device characteristics to determine how the specific properties of a particular device contribute to achieving specific goals, as shown in Fig. 1. We applied two (testing and inquiry) of five method classes (testing, inspection, inquiry, analytical modeling, and simulation) proposed by Ivory and Hearst [10]. Under "testing," a "think-aloud session" was conducted, where an evaluator observed participants' actions (i.e. interacting with the device and performing the task) to determine various usability challenges and witness users' emotional responses (e.g. a decrease in user satisfaction and motivation [20]) and manifested behavioral responses (e.g. an increase in errors, in reactions, or in inefficient or inappropriate activities [20]). Under "inquiry," participants reflected on their emotional and behavioral responses, and the effect those responses had on their use of the devices and associated applications, using a method type "diary". **Participants.** The experiments were carried out in Finland with two age groups (younger than 60 and older than 60) and three different target user groups (students, university staff, fit elderly adults older than 60). Of the sample of 34 participants, 21 were younger or middle-aged, had relatively substantial technological knowledge, and had a positive view of using technology in their daily lives. The second group of 13 were fit elderly participants who were living independently and keen to use new technology to improve their well-being; this group was recruited through direct contact and networking. Members of both groups were at different health-behavior-change stages, as described by the Transtheoretical model (TTM) [18]: pre-contemplation (younger (n = 8), elderly (n = 4)); contemplation (younger (n = 7), elderly (n = 3)); action (younger (n = 1), elderly (n = 2)); and maintenance (younger (n = 5), elderly (n = 4)). The Lappeenranta University of Technology's ethical committee approved the study. All participants were presented with an ethical review statement and informed consent (participants' right to confidentiality, risks, data storage, the use of anonymized data, voluntary participation, no health-related data collected), and a signed consent form was obtained in return. **Procedures and tasks.** In phase 1 (see Fig. 1), we presented the actual demands and resources to the participants, which consisted of:

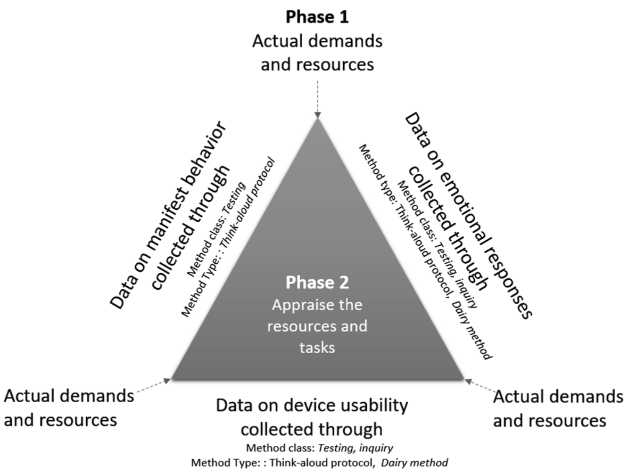


Fig. 1. Methodological approach, image adapted from [28]

- Device presentation: Functioning wearable COTS smartwatches (Apple Watch, Samsung Watch) and pedometers (Misfit Shine 2, Fitbit Charge 2, and Polar A360) were presented to help us to explore the significance of various types of data for future design, as noted by Kanis [11]. These devices were selected based on market availability. No requirements were provided for device selection.
- Timeline: Participants were asked to participate in two one-hour controlled environment sessions (i.e. first meet-up session and final meet-up session), with four weeks of everyday device use between the sessions in a semi-controlled environment.

Experimental tasks: During the first and final meet-up sessions, we assigned experimental tasks (see Appendix A[1]) to be performed to test usability and its effect on participant's emotions and behavior. Usability is one of the most important aspects for the success of any technological product [17], and it has positive correlations with three motivation measures: attention, relevance, and satisfaction [8]; participants' interaction with the device can determine how its specific properties can affect their emotional and behavioral responses. During both sessions, participants were asked to follow a "think-aloud" protocol while performing the presented tasks. In the semi-controlled environment, participants were asked to (i) use devices in real conditions and (ii) complete an open-ended questionnaire in their diaries regarding the devices and associated applications, including any issues they faced or change in the levels of motivation to conduct physical activities, or any other issue they experienced. The aim was to make participants comfortable using the device and to gather data on their emotional and behavioral responses.

In phase 2 (see Fig. 1), consent to collect and use data was presented prior to asking both sets of participants to appraise the resources and tasks set in phase 1. This allowed participants to understand their own perceived demands and resources in using the COTS smartwatches and pedometers. Their appraisal reflected the possible degradation of pathways (i.e. emotional responses and behavioral responses). The effects of these responses on the interactions with the devices were gathered from the participants through diary entries. From this data, we identified commonalities and differences in terms of external and internal contexts.

4 Results

In this section, we synthesize the findings and emphasize commonalities and differences, particularly regarding internal and external contexts (see[2] for Matrix of Study).

C1 Internal Context

C1.1 learning new behavior. Both target groups had to learn new behaviors, such as remembering to charge the device, which affected their daily use of the device. One younger participant stated, "Remembering to charge the device was an issue. I couldn't wear the device because I forgot to put [it] on to charge." Similarly, an elderly participant said, "I didn't put the watch on in the morning, since I took a shower. After that, I forgot to put it on completely." Some general confusion occurred during the evaluation among elderly users when they had to switch between using the external devices and the smartwatches and pedometers. One elderly participant with a Fitbit Charge 2 reported, "Why can't I see my sleep data on the device, while I can see it on my smartphone?" Similarly, another elderly participant noted, "I really can't remember which data I can see on my pedometer and on my mobile phone." However, the younger participants made no such comments.

[1] https://doi.org/10.5281/zenodo.832159.

[2] https://doi.org/10.5281/zenodo.832167.

C 1.2 Meaning of technology and its usefulness. During the first activity, it was surprising to see (i) color of the device and design and (ii) "sleep," "number of steps," and "calories burnt" data being more important than other pedometer/smartwatch functionalities among elderly participants. For example, one elderly participant stated, "I am so excited to see how much I walk a day." Another said, "I just need the band that measures my sleep." However, younger participants placed importance on advanced functionalities, such as receiving calls and texts and the ability to use various applications. One young participant stated, "I would like to have the smartwatch because I want to receive calls." Similarly, during the think-aloud session, in the midst of a lively discussion about privacy invasion by smartwatches and pedometers, there was a positive reaction from both elderly and younger participants regarding how health and physical activity-related data is collected, stored, and analyzed by pedometers and smartwatches. These findings illustrate that the elderly ascribe different meanings to technology than their young counterparts who grew up in a more technological environment [16]. We also found that participants from both groups formed favorable attitudes toward the technology if the devices were useful and relatively easy to utilize.

C 1.3 Transformation in motivation. Some young and elderly participants reported a decrease in motivation after a week of device usage, particularly non-physically active participants who were not willing to engage through data, content, and gamification. In addition, some participants lost motivation due to usability challenges. Indeed, most of the participants in the pre-contemplation or contemplation stage felt that the content did not motivate them, as highlighted in the following: "I see the same information every day; it didn't motivate me to be more physically active." However, participants in the action or maintenance stage [18] engaged though data, content, and gamification; one stated, "The number of goals that have to be achieved motivated me to take more steps."

C 1.4 Transformation of perception towards device characteristics. It was astonishing that, in both groups, the participants' requirements regarding the devices' color and design changed within a week of using them. For example, one elderly participant stated, "I don't like to wear it anymore, because it's white in color and doesn't match my outfit." Conversely, one participant noted, "This color is perfect for me." One younger participant stated, "I have to be very careful when I wear this device, because it's too big," while another younger participant stated, "I can't go to sleep wearing this smartwatch; it's irritating." However, the same participants stated, when selecting the devices, that they looked nice. This change during transformation from the experimentation to habit stage reflects this statement from a previous study [23]: "Doing *something once was an experiment, doing it every day for a week was a habit, and doing it every day for a month was a lifestyle. When attempting to take some new action on a regular basis, one is confronted with many different aspects of the change —how it makes one feel over time" (p. 131).*

C 1.5 Cognitive effort. Our findings revealed that previous knowledge of technological devices (e.g. computers or smartphones) does not decrease the cognitive effort required by the elderly in adapting to new devices. For example, there was an increase in cognitive needs while interacting with smartwatches and pedometers for the first time, and while conducting tasks such as account registration and connecting the pedometers and smartwatches to external devices. Further, increased cognitive effort led to frustration among participants. The participants stated, “I got this device but I don’t know where to start.” Similarly, two other participants said, “It says I have to first register my device, how do I do it?” and “I don’t have an email address, how can I use this?” Another elderly participant commented, “There are so many details to be filled.” Following the elderly users’ frustration, moderators carried out activities such as application installation on external devices, account registration, and connecting the device to Bluetooth.

Another striking observation of cognitive effort requirements occurred while restoring the device. When elderly participants were asked to restore the device during the think-aloud session (i.e. while returning the loaned device), they were unable to do so because of difficulties with navigation or the need for smartphones or computers. One elderly participant remarked, “I cannot find it on my Fitbit; it’s too confusing, do I really have to do this?” This result matched observations from a previous study [26], stating that “the ongoing advance of technology suggests that younger people’s experience with computers will not be a crucial advantage when they grow older.” Young participants also required greater cognitive effort while restoring device. One responded, “It looks like I need my phone to reset my device, which I forgot to bring.” Similarly, another participant commented, “I cannot remove this device from my account using [my] phone; it seems I have to download [an] application on my computer and do it manually.” Participants explained that they lacked practice in restoring devices, and did not have proper instructions for how to do so from the device manufacturer. It seems cognitive effort may occur among younger participants when complex tasks, coupled with a lack of information, are introduced to their busy life schedules.

C2 External Context

C 2.1 Engaging factors. We found that the number of steps taken and data on exercise, heartbeats, calories burned, and sleep statistics were engaging factors for both young and elderly participants. Communication tools such as Skype, Slack, and Telegram were also engaging factors for younger participants.

C 2.2 Device Usability. The COTS smartwatches and pedometers used during the evaluation could be worn on wrists, necks, or ankles; thus, these devices were in close proximity to the bodies [22] of all participants. However, they reported that the device interactions did not satisfy their body shape, size, ability, and dimensions, nor their preferences, interests, and wishes [22]. *The subsequent section describes some commonly reported commonalities and differences in usability factors.*

Font size. Elderly participants complained that the text size on COTS devices with touchscreens was too small to read, stating for example, “I cannot read the text with my reading glasses, can I make this font larger?.”

Interaction with touch screen. During the think-aloud sessions, some elderly participants had difficulties using the touchscreen on pedometers, smartwatches, and external devices due to dexterity problems. In addition, scrolling and navigating within the applications proved difficult for elderly participants. For example, a participant using a pedometer asked, “I pressed the screen on the device but it doesn’t respond; is this device broken?” Another participant who used the smartwatch said that the “touch-screen reacts so fast when I press on it.”

Interaction techniques. During the evaluation, elderly and younger participants regarded the push notifications and reminders differently. A younger participant reported, “I like the device because I could receive all notifications about calls and text data on my watch; I don’t have to use my phone all the time.” This comment reflected a statement from a previous study [14]: “Reminders are the most effective when delivered at the right location, at the right time and the right devices.” However, an elderly participant stated, “Having all the notifications on my watch with vibration feels so irritating and like getting an electric shock.” While the young group of participants found receiving notifications and reminders through COTS devices useful, their counterparts felt the opposite, which is in line with a previous study’s finding [14] that “age might however influence the interaction techniques.”

Reliability and accuracy. The data’s reliability was a concern for both groups of participants. For example, one of the elderly participants reported, “It didn’t record one of my afternoon naps. How can I rely on the sleep analysis data?” Similarly, a younger participant stated, “I had the device with me when I went to the fitness center, but there was no change in fitness activities.” According to another younger participant, “Sometimes I feel the measuring data isn’t accurate. For example, I was sitting and working, but the app shows I am resting.”

Device connectivity. Connecting the wearables and the external device, and synchronizing the data using Bluetooth technology, were the most commonly reported usability challenges by both groups of participants. For example, one younger participant stated, “Connecting the phone with the watch, I had to turn on and off the Bluetooth all the time.” An elderly participant reported, “I got an error on my application. My Charge 2 isn’t syncing because my phone’s Bluetooth is off, but the Bluetooth on my phone is on.”

Battery: Both older and younger participants raised concerns about the battery. As one of the younger participants reported, “Using the watch is easy, but keeping track of the battery is a problem.” Another participant stated, “The battery runs out quickly.” Participants with an integrated battery (e.g. Misfit COTS pedometers) had usability advantages over the other smartwatches and pedometers, as there were no comments regarding battery issues. Elderly users reported that it was difficult to parallel the use of the application installed on the external devices and COTS pedometers without any display. One participant commented, “When I was walking and wanted to see how long I had walked, it was difficult to take out the phone and view data.”

5 Discussion and Conclusion

Here, we will discuss the results of the evaluation of both elderly and younger participants, present implications for practice, and reveal our research findings. In addition, we will offer suggestions for future work. This study involved a small number of participants in a limited geographic location, meaning the generalizability of the results may not be possible; thus, all stakeholders, including device manufacturers and application developers, should take the findings as suggestions rather than conclusive evidence.

The first finding showed that both the internal and external contexts had a cause-effect relationship with both target groups, with more commonalities than differences in terms of the internal context, especially regarding usability factors of the external context and the users' own perceptions of the devices. Therefore, it would be beneficial to integrate both contexts during the design of wearable devices and their associated applications. The data gathered from emotional responses and manifested behavior showed that the internal context can strongly influence any age group if the effect on the external context appears or vice-versa; it can obstruct the acceptance of COTS smartwatches and pedometers by changing an individual's motivation. Further, the higher the degree of external context (i.e. usability factors), the better the internal context.

The most common external context usability elements that affected the use of wearable devices included font size, interaction with the touchscreen, interaction techniques, and applications installed; these strongly influenced age-related deficits and are in line with previous studies [1]. Device connectivity, battery life, reliability, and accuracy were the most commonly cited common important internal factors, which also aligns with previous studies [27]. Further, these results may change, depending on the context in which individuals use COTS devices. Future studies should measure how quickly both the internal and external contexts that can obstruct device usage appear in large numbers within both target groups over a specified period, and both elderly and younger individuals could retain the COTS device after appearance of cause-effect relationships.

Interestingly, despite having all user engagement features, such as data, gamification, and content, on either wearable devices or external devices with associated applications, these extrinsic motivational factors did not have a long-term effect on physically inactive participants who were in either the pre-contemplation or contemplation stage. Hence, for a person to be physically active, intrinsic motivation must evolve on its own, while extrinsic motivation will only enhance intrinsic motivation. Further studies can implement self-determination theory to discover which influential factors might awaken the intrinsic motivation of individuals in the pre-contemplation or contemplation stages of behavior change. First impressions of the devices were temporary for both groups, which likely faded based on the individuals' context of use and hierarchy of needs, whether cognitive or psychological. This finding led us to understand that a changed impression might affect the motivation to use the wearable device long term. Therefore, future work could develop guidelines that include the hierarchy of needs of both younger and elderly individuals based on the context of use of COTS wearable devices, which could help device manufacturers and application developers create sustainable COTS devices and associated applications.

To understand the commonalities and differences between younger and elderly participants using the same COTS devices, we developed experimental tasks. The results found commonalities in terms of internal context in both participant groups, apparently due to both the effect and usefulness of the external context. Therefore, certain measures should be taken regarding the external context, such as including age-appropriate smartwatch and pedometer device characteristics to reduce the cause-effect relationship of the internal and external contexts. Users will then feel comfortable and develop a high degree of satisfaction, motivation, and enjoyment regarding these devices' usefulness. The new design could decrease manifested negative behaviors and emotional responses by increasing the acceptance of COTS smartwatches and pedometers. For the elderly, appropriate font sizes and better interaction with the touchscreen and associated applications, as based on their hierarchy of needs, could improve their manifested behaviors and emotional responses and increase their satisfaction, leading to them adopting the devices for longer. Our future work will investigate: (i) how the internal and external contexts differ when secondary users, such as caregivers or relatives, use COTS smartwatches and pedometers on behalf of frail elderly users; (ii) the strong bond between the two contexts through an empirical study; and (iii) differences caused by geographical area, gender, and/or culture when repeating the same study with a larger sample of participants.

Acknowledgement. We would like to thank Miina Sillanpää Foundation, LUT Research Platform on Smart Services for Digitalization (DIGI-USER) for their generous support of our research and Greg Priest-Dorman for valuable feedback on early version of the paper.

References

1. Angelini, L., et al.: Designing a desirable smart bracelet for older adults. In: UbiComp 2013 Adjunct, pp. 425–434 (2013)
2. Asimakopoulos, S., et al.: Motivation and user engagement in fitness tracking: heuristics for mobile healthcare wearables. In: Informatics, vol. 4, no. 1, p. 5 (2017)
3. Charness, N., et al.: Aging and information technology use: potential and barriers. Curr. Dir. Psychol. Sci. **18**(5), 253–258 (2009)
4. Chodzko-Zajko, W.J., et al.: Exercise and physical activity for older adults. Med. Sci. Sports Exerc. **41**(7), 1510–1530 (2009)
5. Fukuda, R., Bubb, H.: Eye tracking study on web-use: comparison between younger and elderly users in case of search task with electronic timetable service. PsychNology J. **1**(3), 202–228 (2003)
6. Gregor, P., et al.: Designing for dynamic diversity: interfaces for older people, Edinburgh, May 2002
7. Gwizdka, J.: What's in the context? In: CHI 2000 Work (2000)
8. Hu, Y.: Motivation, usability and their interrelationships in a self-paced online learning environment. ProQuest Diss. Theses, 158 p. (2008)
9. International Organization for Standardization: ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) - part 11: guidance on usability. International Organization for Standardization, vol. 1998, no. 2, p. 28 (1998)

10. Ivory, M.Y., Hearst, M.A.: The state of the art in automating usability evaluation of user interfaces. ACM Comput. Surv. **33**(4), 470–516 (2001)
11. Kanis, H.: Usage centred research for everyday product design. Appl. Ergon. **29**(1), 75–82 (1998)
12. Khakurel, J., et al.: Tapping into wearable devices revolution for enterprise use. Inf. Technol. People 1–17 (2017)
13. Leonardi, C., et al.: Designing a familiar technology for elderly people. Gerontechnology **7** (2), 151 (2008)
14. McGee-Lennon, M.R., et al.: User-centred multimodal reminders for assistive living. In: Proceedings of 2011 Annual Conference on Human Factors in Computing Systems, CHI 2011, pp. 2105–2114 (2011)
15. Nelson, M.E., et al.: Physical activity and public health in older adults: Recommendation from the American College of Sports Medicine and the American Heart Association. Circulation **116**(9), 1094–1105 (2007)
16. Oksman, V.: Young people and seniors in finnish “mobile information society”. J. Interact. Media Educ. **2006**, 1–21 (2006)
17. Paz, F., Pow-Sang, J.A.: A systematic mapping review of usability evaluation methods for software development process. Int. J. Softw. Eng. Appl. **10**(1), 165–178 (2016)
18. Prochaska, J.O., Velicer, W.F.: The transtheoretical model of health behavior change (1997). http://www.ncbi.nlm.nih.gov/pubmed/10170434
19. Rasche, P., et al.: Activity tracker and elderly. In: 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, pp. 1411–1416 (2015)
20. Robertson, S.A.: Contemporary Ergonomics 1995, pp. 62–69. Taylor & Francis, Milton Park (1995). https://books.google.fi/books?isbn=0748403280
21. Siek, K.A., Rogers, Y., Connelly, K.H.: Fat finger worries: how older and younger users physically interact with PDAs. In: Costabile, M.F., Paternò, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 267–280. Springer, Heidelberg (2005). https://doi.org/10.1007/11555261_24
22. Suh, A., et al.: The use of wearable technologies and body awareness: a body–tool relationship perspective. In: HCI International 2016, pp. 388–392 (2016)
23. Tomlinson, B.: Greening Through IT, Information Technology for Environmental Sustainability. The MIT Press, Cambridge (2010)
24. Tudor-Locke, C., et al.: How many steps/day are enough? For older adults and special populations. Int. J. Behav. Nutr. Phys. Act. **8**(1), 80 (2011)
25. U.S. Department of Health and Human Services: 2008 Physical activity guidelines for Americans. Pres. Counc. Phys. Fit. Sport. Res. Dig. **9**(4), 1–8 (2008)
26. Van de Watering, M.: The impact of computer technology on the elderly. http://www.marekvandewatering.com/texts/HCI_Essay_Marek_van_de_Watering.pdf
27. Young-Hyun, P., Kwang-hee, H.: Information display of wearable devices through sound feedback of wearable computing. In: Jacko, J.A. (ed.) HCI 2007. LNCS, vol. 4551, pp. 1200–1209. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73107-8_132
28. Young, R.A., et al.: The action-project method in counseling psychology. J. Couns. Psychol. **52**, 2 (2005)
29. Young, R.A., Valach, L.: Action and language: contextual action theory in counselling. Psychol. Française **61**(1), 31–42 (2016)
30. Zhou, X., et al.: Assessing age-related performance decrements in user interface tasks. In: 2011 IEEE ICIA 2011, pp. 817–822 (2011)

mHealth Platform for the Delivery of Rehabilitation and Physical Exercise at Home for Parkinson's Disease Patients

Fernando Martínez-Martí, Samanta Villanueva-Mascato(✉), Benjamin Belletoile, Jorge Cancela, and María Teresa Arredondo

Life Supporting Technologies (LifeSTech),
Universidad Politécnica de Madrid (UPM), Madrid, Spain
{fmartinez,svillanueva}@lst.tfo.upm.es
http://www.lst.tfo.upm.es

Abstract. Parkinson's disease (PD) is a neurodegenerative and progressive disorder of the central nervous system that affects mainly the motor system. As a consequence of the disease, PD patients suffer a progressive reduction of their independence and Quality of Life (QoL). Literature research shows that the use of auditory, visual and haptic cueing could significantly benefit the motor performance of PD patients. Likewise, music therapy (MT) has shown notable benefits for PD either in the motor and non-motor dimensions. This work is aimed at designing a mHealth system to deliver a home based rehabilitation programme for PD patients based on external cueing and MT called HOOP, integrated by a set of inertial sensors, an Android application, and a web tool for professionals. This designing process has been performed according to expert's consultation in order to develop a tool that can cover both patients and professional necessities.

Keywords: Parkinson · Rehabilitation · Remote training
Rhythmic auditory stimulation · Music therapy

1 Introduction

Parkinson's disease (PD) is one of the most common neurodegenerative disorders affecting more than 1% of people older than 60 years (and with an increased prevalence in older groups). Progression of the disease is strongly correlated with higher costs – both for patients and healthcare institutions – and, even if a cure is not yet available, considerable savings can be achieved by slowing down this progression [1]. Many recent studies have shown that main motor symptoms of Parkinson's disease (bradykinesia, hypokinesia, restring tremor, rigidity, and postural instability) have an important impact in poorer Quality of Life (QoL) of PD patients, both in the mobility sub-scale of PDQ-39 and overall QoL in PD [2–4].

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 42–50, 2018.
https://doi.org/10.1007/978-3-319-76111-4_5

In this sense, these motor abnormalities may be improved by the use of stimulation techniques that help to perform training and rehabilitation exercises such as listening to marching music. These types of techniques are called Rhythmic Auditory Stimulation (RAS) [5–7]. There are several approaches about how to use RAS and music with PD patients and healthy patients (over 65), this type of experiments are usually based on recognition of gait-health related problems [8]. Styns et al. detected that healthy young adults walked faster with music than with metronome cues [9], while Wittwer et al. found that healthy adults over 65 increased their cadence with both music and metronome, but stride length and gait velocity was increase only with music [10]. Others experiments have tested this stimulation with PD patients obtaining promising results [11].

However, most of the research works only focus on the stimulation techniques and put no effort in providing therapies to patients and combining these rehabilitation sessions within RAS techniques. Furthermore, sometimes this system does not evaluate if training therapies are done correctly. Many studies have proved the efficacy of including new technologies to promote active ageing integrating external sensors [12] and mobile applications [12,13]. This work presents HOOP, that is a platform designed to provide rehabilitation and training sessions for PD patients at home evaluating its performance, with a constant supervision of therapists and clinicians through a web system. In addition, therapist/clinician will be able to modify and personalise the training and rehabilitation sessions according to users performance, providing them a personalised therapist which will increase engagement and motivation.

2 Materials and Methods

2.1 Overall Architecture

The architecture presented in this work is based on three different modules which work together to offer an integrated mHealth tool for Parkinson's disease patients and caregivers to provide a more optimal training and rehabilitation therapies, and to be able to evaluate their performance (see Fig. 1): HOOP sensors, HOOP Therapist Website, and HOOP for Mobile.

2.2 Designing Procedure

HOOP has been designed according to a set of interviews that were carried to several participants to define requirements and needs of this mHealth tool. A total of 12 participants (4 males and 8 females) took part of the study covering different professional profiles: four therapists (three of them with experience on music therapy for PD), two occupational therapists, two neurologists, one physiologist and three biomedical engineers, all of them familiar with the use of External Rhythmical Cueing (ERC) and Music Therapy (MT) on PD. The

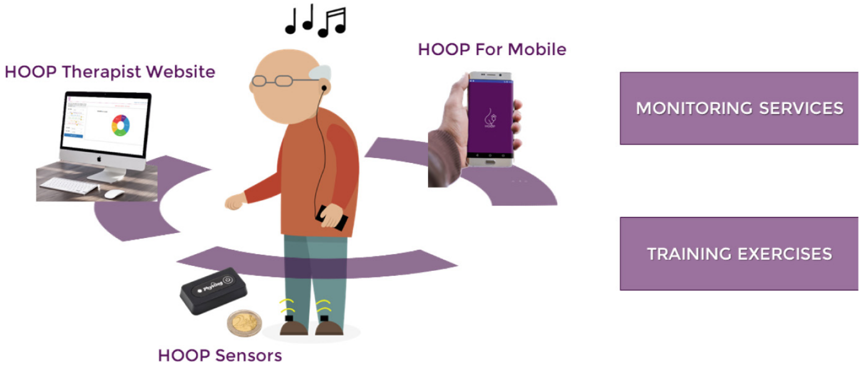


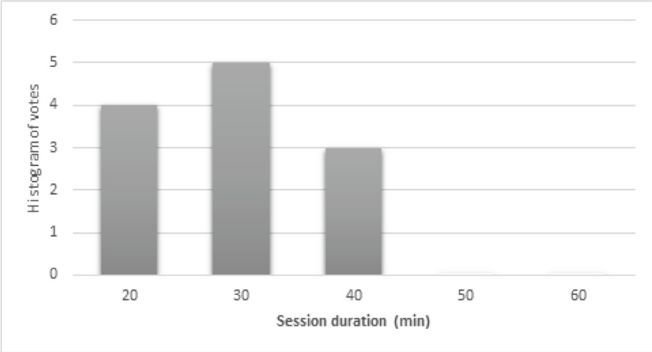
Fig. 1. HOOP architecture.

average age of the group was 35.25±9.44 years old and they presented an average of 7.33 ± 4.88 years of experience with PD patients.

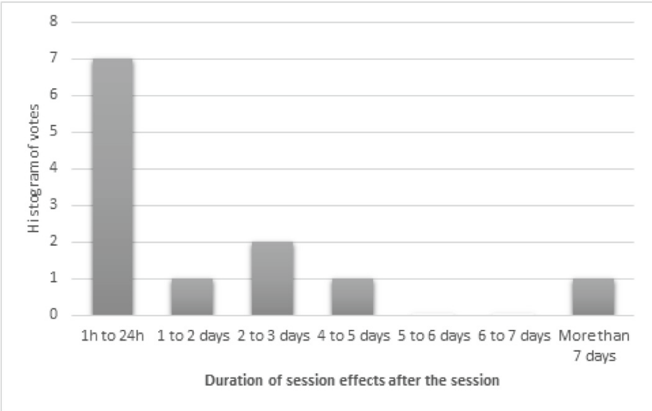
At first, experts were inquired with a set of questions regarding general aspects of how a rehabilitation programme at home for PD patients should look like. This set of questions help to identify the optimal duration for a daily gait rehabilitation sessions (Fig. 2(a)), how much time to the effects of auditory cueing and/or music therapy last after the session (Fig. 2(b)), and how to increase adherence of the patients to the rehabilitation methods (Fig. 2(c)). The aim of these questions was to adapt the development according to experts answers.

Furthermore, experts were asked (see Fig. 3) about the User Interfaces (UI) for both the patient (HOOP for Mobile) and the professional application (HOOP Therapist Website). For the patient side, a prototype of the application has been designed in Android Environment. For the assessment of the UI a variation of the System Usability Scale (SUS) [14] was used, although in this case questions were adapted to ask the experts whether they think that the patients would be able to deal with the app. The aim of this part of the work was to explore which were the therapists' needs in this scenario.

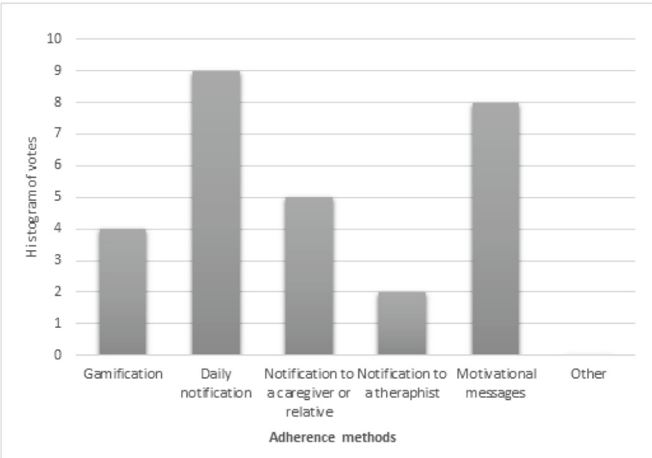
In a previous work multiple mock-ups of auditory cues were designed and explored to find out the characteristics of a Rhythmic Auditory Stimulation to be used for such purpose [15]. As a result, several aspects regarding the rhythmic, melodic and harmonic components of the auditory cues were identified in order to create new stimuli to facilitate and engage the performance of training and rehabilitation tasks.



(a)



(b)



(c)

Fig. 2. Results from questionnaires with the experts. Recommended session duration (a), duration of effects after the session (b), and the best methods to ensure adherence (c)

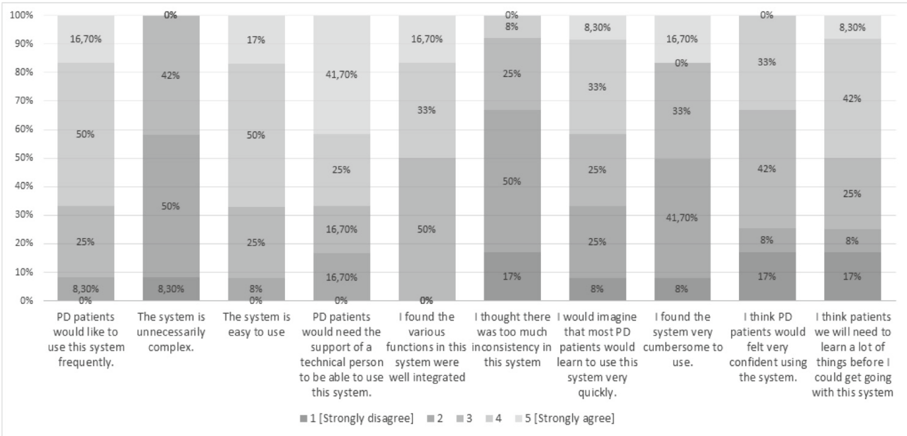


Fig. 3. Responses to the experts' opinion about the usability of the application (SUS test). Each question presents the results grouped into categories according to the experts' opinions from 1 (Strong disagree) to 5 (Strong agree).

3 Results

3.1 Sensor System

Current version of HOOP uses a set of commercial inertial sensors. These sensors are Physilog 5 (GaitUp^{TM}, Switzerland). They consist of a three axis gyroscope and three axis accelerometer, with a low energy bluetooth link, and a μSD memory card.

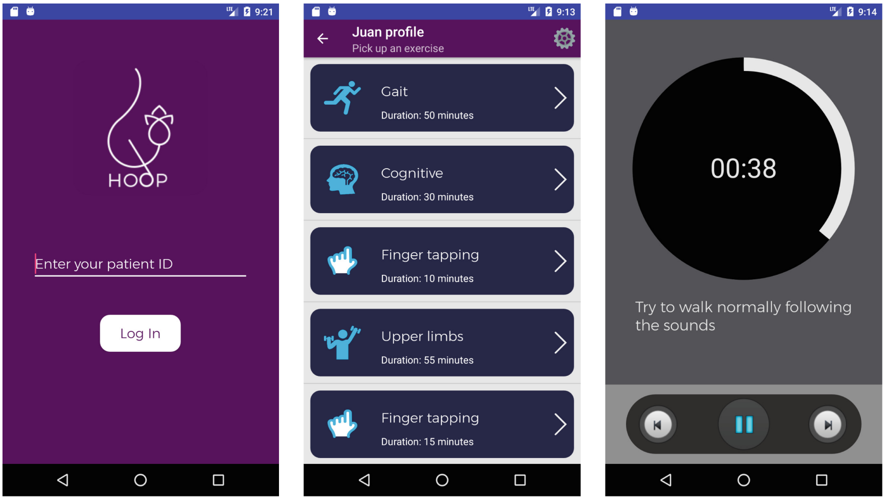
(a) (b)

Fig. 4. HOOP sensors placement. In the wrist (a) during the upper limb training sessions and in the ankle for lower limb (b).

The use of these sensors will allow to evaluate the performance of the exercises by the clinician or therapist. They will be measure inertial movements and compared with the exercises and music performance. Sensors will be used in pairs to distinguish left and right arm/foot. They will be located at the wrists while evaluating upper limbs exercises and in the ankles for gait training sessions (see Fig. 4).

3.2 Mobile Application

HOOP for Mobile consists in an Android Application that interacts with the Parkinson's patients and establishes a set of exercises according to the scheduled activities by the therapists and clinicians. It also will collect information from inertial sensors connected by Bluetooth in order to evaluate the correct performance of the exercises by evaluating the movement of the arms in upper limbs exercises or the steps in gait exercises. During exercises both music and cueing stimulation are provided to the patients in order to stimulate and help them to maintain a cadence performance.



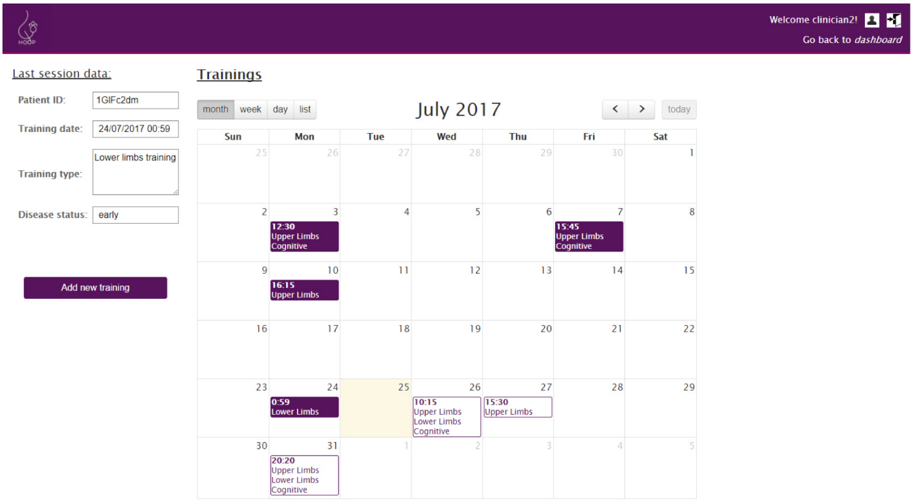
(a) Login (b) Scheduled exercises (c) Gait exercise

Fig. 5. HOOP for Mobile snapshots. Login activity (a) to identify the patient, the set of exercises (b) previously scheduled by the therapist, and the performance of a gait exercise (c).

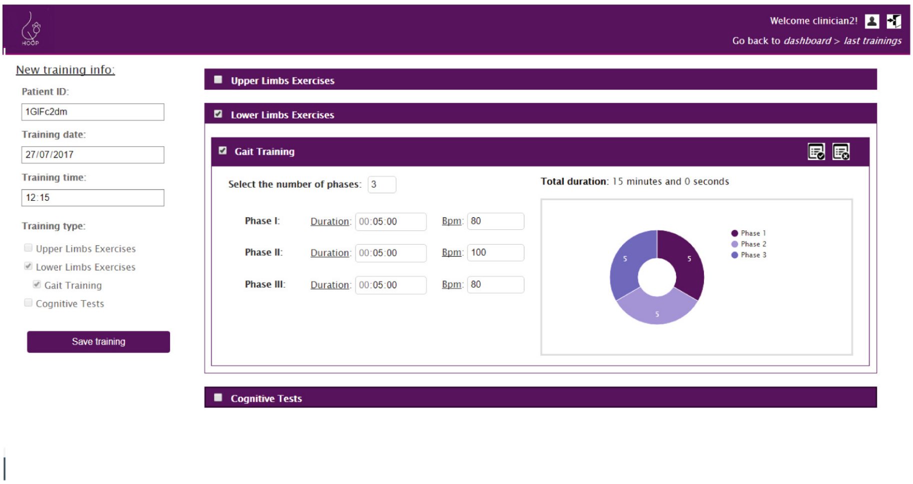
Figure 5 shows some snapshots from HOOP for Mobile application. Once the resulting information from the exercises performance is collected, it is updated in the database to allow clinician/therapist to evaluate the outcomes achieved by the patient and to establish changes in the therapy according to the patient needs, providing a more personalised therapy.

3.3 Therapist Website

The main objective of the HOOP Therapist Website is to provide a friendly user interface for the Therapists and Clinicians in order to ease their supervision tasks on those Parkinson patients who perform rehabilitation and training exercises



(a) Calendar view



(b) Create new training session

Fig. 6. HOOP Therapist Website snapshots. Calendar view (a) in which the therapist/clinician can set or edit exercises and view the results from previous training sessions. The definition of a new training session (b) which will be scheduled in the patient exercises list.

under their care. HOOP Therapist Website is implemented using Django (version 1.11.1) on top of Python (version 3.4.3) as backend, while HTML, Javascript, and CSS are used for frontend.

Every therapist or clinician is able to evaluate the performance of previous sessions (Fig. 6(a)) using the calendar view. By pressing in one of past training dates, a set of plots which are used to analyse the outcomes from the exercises allows the professional to set a new interventions by creating a new personalised set of exercises in an specific date Fig. 6(b).

4 Conclusions

According to the successful outcomes obtained for those studies which combine gait or physical therapies with auditory cueing or music stimulation, we have developed an integrated platform that eases the performance of rehabilitation and training sessions of Parkinson's disease patients at home. It is integrated by three modules: HOOP Sensors, based on a set of commercial inertial sensors, which are used to evaluate exercise performance together with mobile application; HOOP for Mobile, that is an Android application, which interacts directly with the patients, presents their scheduled activities and collects information from sensors. And finally the HOOP Therapist Website, that is a webtool for clinicians and therapists oriented to provide to professionals with tools which help them to evaluate the performance of the training sessions and to set a personalised set of exercises for Parkinson's disease patients with the goal to reduce the impact of motor symptoms and them, improve their quality of life.

The design and development of HOOP have been performed according to experts opinion. A set of questions has been given to compile information about what are the needs of typical rehabilitation and training sessions for Parkinson's disease patients to adapt the architecture and design to these requirements. And once the application was done, an usability test was prepared to evaluate the opinion from these groups of experts not only about patient application but also clinician web application. Obtaining in both case very good results that shows promising results for the validation with PD patients.

Acknowledgement. Authors would like to thank mHealth tOol for parkinsOn's disease training and rehabilitation at Patient's home (HOOP) project, funded by EIT-Health consortium under Innovation by Ideas initiative, and to PD_manager project, funded within EU Framework Programme for Research and Innovation Horizon 2020.

References

1. Lindgren, P., von Campenhausen, S., Spottke, E.A., Siebert, U., Dodel, R.: Cost of Parkinson's disease in Europe. Eur. J. Neurol. **12**(s1), 68–73 (2005)
2. Peto, V., Jenkinson, C., Fitzpatrick, R., Greenhall, R.: The development and validation of a short measure of functioning and well being for individuals with Parkinson's disease. Qual. Life Res. **4**(3), 241–248 (1995)

3. Schrag, A., Selai, C., Jahanshahi, M., Quinn, N.P.: The EQ-5D-a generic quality of life measure-is a useful instrument to measure quality of life in patients with Parkinson's disease. J. Neurol. Neurosurg. Psychiatry **69**(1), 67–73 (2000)
4. Diamond, A., Jankovic, J.: The effect of deep brain stimulation on quality of life in movement disorders. J. Neurol. Neurosurg. Psychiatry **76**(9), 1188–1193 (2005)
5. de Dreu, M.J., van der Wilk, A.S.D., Poppe, E., Kwakkel, G., van Wegen, E.E.H.: Rehabilitation, exercise therapy and music in patients with Parkinson's disease: a meta-analysis of the effects of music-based movement therapy on walking ability, balance and quality of life. Parkinsonism Relat. Disord. **18**, S114–S119 (2012)
6. Patel, N., Jankovic, J., Hallett, M.: Sensory aspects of movement disorders. Lancet Neurol. **13**(1), 100–112 (2014)
7. Paulson, S., Bharucha, J., Iyer, V., Limb, C., Tomaino, C.: Music and the mind: the magical power of sound. Ann. N. Y. Acad. Sci. **1303**(1), 63–79 (2013)
8. Pogorelc, B., Bosnić, Z., Gams, M.: Automatic recognition of gait-related health problems in the elderly using machine learning. Multimedia Tools Appl. **58**(2), 333–354 (2012)
9. Styns, F., van Noorden, L., Moelants, D., Leman, M.: Walking on music. Hum. Mov. Sci. **26**(5), 769–785 (2007)
10. Wittwer, J.E., Webster, K.E., Hill, K.: Music and metronome cues produce different effects on gait spatiotemporal measures but not gait variability in healthy older adults. Gait Posture **37**(2), 219–222 (2013)
11. Lopez, W.O.C., Higuera, C.A.E., Fonoff, E.T., de Oliveira Souza, C., Albicker, U., Martinez, J.A.E.: Listenmee® and Listenmee® smartphone application: synchronizing walking to rhythmic auditory cues to improve gait in Parkinson's disease. Hum. Mov. Sci. **37**, 147–156 (2014)
12. Sendra, S., Granell, E., Lloret, J., Rodrigues, J.J.P.C.: Smart collaborative mobile system for taking care of disabled and elderly people. Mob. Netw. Appl. **19**(3), 287–302 (2014)
13. Lane, N.D., Lin, M., Mohammod, M., Yang, X., Lu, H., Cardone, G., Ali, S., Doryab, A., Berke, E., Campbell, A.T., Choudhury, T.: BeWell: sensing sleep, physical activities and social interactions to promote wellbeing. Mob. Netw. Appl. **19**(3), 345–359 (2014)
14. Jordan, P.W.: Usability Evaluation in Industry. Taylor & Francis, Milton Park (1996)
15. Cancela, J., Moreno, E.M., Arredondo, M.T., Bonato, P.: Designing auditory cues for Parkinson's disease gait rehabilitation. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5852–5855. IEEE, August 2014

Context-Aware Recommendations for Sustainable Wardrobes

Anders Kolstad[1,2](✉), Özlem Özgöbek[1], Jon Atle Gulla[1], and Simon Litlehamar[2]

[1] Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway
{ozlem.ozgobek,jon.atle.gulla}@ntnu.no
[2] Accenture AS, Oslo, Norway
{anders.kolstad,simon.litlehamar}@accenture.com

Abstract. Through recycling textile waste, greenhouse gas emissions can drastically be reduced. Such textile recycling has become a lot easier with clothing retailers now starting to offer recycling checkpoints. Moreover, people today are often challenged by overloaded wardrobes and store many clothing items that they never use. In this paper, we describe an Internet of Things system that creates incentives for the users to recycle their clothes, benefiting the environmental sustainability. We propose a content-based recommendation approach that utilizes semantic web technologies and that leverages a set of context signals obtained from the system's architecture, to recommend clothing items that might be relevant for the user to recycle. Experiments on a real-world dataset show that our proposed approach outperforms a baseline which does not utilize semantic web technologies.

Keywords: Internet of Things · Recommender systems
Content-based recommendation · Textile recycling · Linked open data
Bag of concepts

1 Introduction

In today's world, overconsumption is becoming a huge concern. The equivalent for fast food called 'fast fashion' is becoming a phenomenon, and people continue to purchase large quantities of clothes. Because of this, textile recycling has become a pressing issue. Textile recycling has old history but was earlier often just concerned with the economic benefits. As of now, the environment has become the number one factor for doing textile recycling. For every pound of recycled textile, more greenhouse gas emissions are prevented than for every pound of glass, plastic, and paper—combined [1]. Moreover, Klepp and Laitala found that 20% of the clothes in Norwegians' wardrobes were never or rarely used [2]. Because of this, clothing retailers are now starting to offer checkpoints where people can recycle their clothes in order to enhance the environmental sustainability.

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 51–60, 2018.
https://doi.org/10.1007/978-3-319-76111-4_6

It is important that people get incentives to use these recycling checkpoints. Such incentives can be generated by recommender systems and Internet of Things technology. Traditionally, recommender systems try to predict items that might be of interest to the users. A popular technique for recommender systems is known as *content-based filtering* [3]. Content-based recommendations build a user profile of item properties that the user has shown an interest of in the past and then computes item similarities with other items that the user has not seen yet. In previous studies, content-based recommender systems enabled with semantic web technologies have shown promising results by increasing the accuracy of the recommendations [4]. Linked Open Data (LOD) is a semantic web technology that forms a set of rules for publishing data so that the data become machine-readable and free of use for anyone [5].

In [6], we proposed a system called Connected Closet—a smart closet where clothing items enabled with radio-frequency identification (RFID) tags can be scanned in the user's closet, tracking the usage history of the clothes. The system's mobile application leverages data collected from the smart closet, to guide users in their everyday lives and making the users' wardrobes explorable on a mobile device. In [7], we proposed a recommendation approach for recommending daily outfits to the users. In this paper, we propose a semantic content-based recommender system that leverages a set of context signals obtained from the system's architecture to provide recycling suggestions to users of the system. By utilizing semantic web technology, the recommender system's accuracy improves. Moreover, it can improve the system's transparency and increase the user's trust and confidence in the system.

Contributions. The main contributions of this paper may be summarized as follows:

1. We propose a content-based recommender system that utilizes LOD to recommend clothing items to be recycled from a smart closet.
2. We evaluate our proposed recommender system on a real-world dataset and compare it to a baseline that does not utilize semantic web technology.

The remainder of this paper is structured as follows. In Sect. 2, we describe the contextual models addressed when computing the recommendations. Then, we describe our recommendation algorithm in Sect. 3. In Sect. 4, we evaluate our proposed approach and discuss related work in Sect. 5. We conclude with a summary and discuss future work in Sect. 6.

2 Contextual Models

In this section, we describe the contextual signals obtained from the proposed system's architecture depicted in Fig. 1. We describe how they are obtained and how they affect the recommendation algorithm for recycling recommendations.

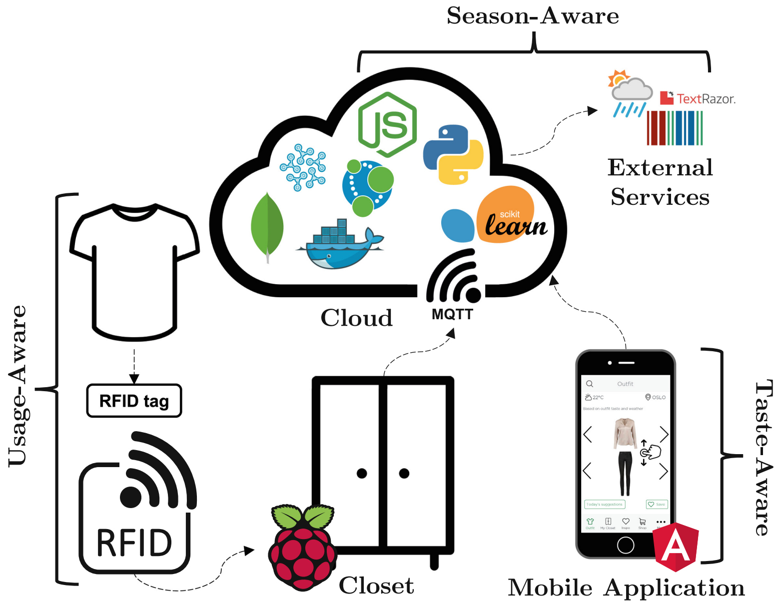


Fig. 1. Architecture for a sustainable wardrobe. The figure indicates where in the system the contextual signals are obtained from.

2.1 Usage-Aware

The user's clothing items are enabled with RFID tags. These can be manually scanned through an RFID reader connected to a tiny computer embedded in the user's Closet. The computer will broadcast a message containing information about the scan to a set of services deployed in the Cloud. Each scan is added to the set S.

Clothing items that are often checked out of the user's closet will achieve a higher user rating. More formally, based on item usage, we calculate the user's rating of a clothing item as follows:

$$\hat{r}_{\mu}(u,i) = \frac{1}{2}|S_{u,i}|, \tag{1}$$

where $S_{u,i}$ is the set of all scans of clothing item i done by user u.

2.2 Taste-Aware

In the Mobile Application, the user can save his favorite outfits. An outfit is represented as a tuple (one top and one bottom). The user's favorite outfits are added to the set O_u. Items that occur in many outfit combinations will achieve a high rating. More formally, the user rating based on the user's favorite outfits is defined as follows:

$$\hat{r}_{\tau}(u,i) = |\{(j,k) \in O_u \mid j = i \vee k = i\}| \tag{2}$$

2.3 Season-Aware

The usage pattern of some items might only occur during a season, e.g., a winter coat will only be used in the winter. Such seasonal clothing items are assigned to a season {winter|spring|summer|fall}. If a seasonal clothing item is recommended for recycling, we check if the item was used during the last assigned season. This is done by looking up the latest item scans in S. If the item was used during the season, the item is removed from the recommended list and not displayed to the user.

3 Semantic Content-Based Recommender System

In [6], we proposed a model for recycling recommendations that recommended the lowest rated items. Although, considering how the ratings are obtained from the context signals, newly bought clothing items would always be recommended for recycling when using such a model. In this paper, we propose a content-based recommender system that recommends the items that are as least similar as possible to the user profile.

3.1 Vector Space Model

In our recommendation approach, we adopt the Vector Space Model, where we represent each clothing items as a vector. We then use the *Bag of Concepts* [14] approach to create the vectors using entities from `Wikidata`[1], which is a Knowledge Base published as LOD. As a weighting scheme, the Concept Frequency (CF) is used. Here, the term concept refers to the `Wikidata` entities.

The user profile is represented as a set of clothing item vectors:

$$profile(u) = \{\mathbf{i} \mid \hat{r}(u, i) > \lambda\}, \tag{3}$$

where $\hat{r}(u, i)$ is an aggregate of $\hat{r}_\mu(u, i)$ and $\hat{r}_\tau(u, i)$, and $\mathbf{i}$ is the CF vector representing clothing item i.

The user's clothing items are then sorted ascending in a ranked list using this scoring function:

$$\bar{r}(u, i) = \frac{\sum\limits_{\mathbf{j} \in profile(u)} dist(\mathbf{i}, \mathbf{j})}{|profile(u)|}, \tag{4}$$

where $dist(\mathbf{i}, \mathbf{j})$ is a distance measure between the vectors representing the clothing items i and j. In our approach we use the *Euclidean distance* defined as follows:

$$dist(\mathbf{q}, \mathbf{p}) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}, \tag{5}$$

where $\mathbf{q}$ and $\mathbf{p}$ are both vectors of n dimensions.

[1] https://www.wikidata.org/.

3.2 Semantic Item Representation

Figure 2 shows an excerpt of the services deployed in the Cloud. When new clothing items are fed into the system, they are inputted with a free text description to the Catalog service. To represent the clothing items as vectors using `Wikidata` entities with CF, the following process is performed on the item's text description (the location of where in the system each step is performed is depicted in Fig. 2):

(1) *Entity extraction.* We extract `Wikidata` entities by using the `TextRazor API`[2]. For disambiguating entities, each entity is ranked with a confidence score based on multiple signals in the text.
(2) *Weighting.* We generate vectors using the CF weighting scheme from the entities returned by Step 1.
(3) *Storing.* The vectors are then stored in the system's graph database called Item storage.

In the system's current stage, removal of stop entities and generic entities, is not addressed and will be included in the process in future research.

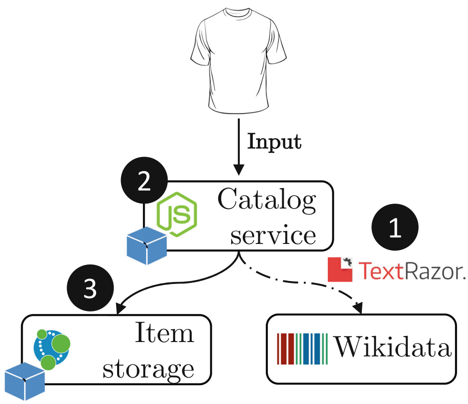


Fig. 2. Excerpt of the system's services deployed in the Cloud. An illustration of the process of a new item being stored in the system.

4 Experiments

To demonstrate the validity of our approach, we perform an evaluation on a dataset collected from the Web.

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

4.1 Dataset

In order to evaluate our approach, we desired a dataset consisting of users with outfit ratings, the user's usage history on clothing items, and information on which clothing items that have been recycled. Because our system is not yet in full scale production and there seems to be little available data in this domain, we could only obtain a dataset addressing outfit ratings.

The dataset was collected from the social media site Polyvore[3]. Polyvore is a site where users can create fashion outfits. Other people can rate these outfits using a 'like' button on a unary rating scale. This mirrors the functionality found in the smart closet's mobile app. The collected dataset consists of 260 rated outfits composed by 158 clothing items, 7093 users and 19287 outfit ratings.

Figure 3 shows an example from one of the items in the collected dataset. The top shows the representation of the item using the classic *Bag of Words* approach with a Term Frequency (TF) vector. Below, is the representation of the item after it has gone through the process described in Sect. 3.2. The blue node (`item`) represent the clothing item, while the green nodes (labels starting with'`Q`') represent the `Wikidata` entities describing the clothing item. To visualize the context signals, we have included one user's interactions with the item. In the figure, $\hat{r}_\mu(u,i)$ (`USAGE_COUNT`) and $\hat{r}_\tau(u,i)$ (`OUTFIT_LIKES`) are represented as relations from the `user` node to the `item` node.

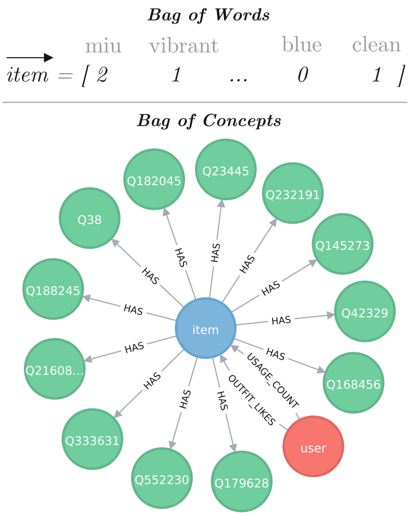


Fig. 3. Representation of a sample item. Text description: *"Miu Miu's vibrant Resort '17 collection is inspired by the '90s rave scene. Knitted in a kaleidoscope of hues, this cropped sweater has sumptuous touches of wool, mohair and [...] Dry clean.. Made in Italy..".* Given this description, the `Wikidata` entities such as `Q42329` (wool), `Q232191` (sweater), and `Q552230` (Miu Miu) are extracted.

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

4.2 Evaluation Method

Due to the nature of the dataset, we can only consider the ratings obtained by the Taste-Aware context signal. For this reason, the other signals are neglected in this experiment. Moreover, the dataset does not contain clothing items that have been recycled by the users. This means that we need to make an assumption for when a clothing item is relevant for recycling. In this experiment, a clothing item that is relevant for recycling is a clothing item that occurs only once in the user's favorite outfits, i.e., $\hat{r}_\tau(u,i) = 1$.

Evaluation protocol. We evaluate all the users in the dataset that has at least one item i such that $\hat{r}_\tau(u,i) > 3$. Moreover, we set $\lambda = 2$ in Eq. 3. We then assume that the user only owns items that occurs at least once in his favorite outfits, i.e., $\hat{r}_\tau(u,i) > 0$. For these items, we generate a recommended list to the users using Eq. 4.

Evaluation metrics. To assess the quality of the recommendations we apply the traditional evaluation metrics Recall and Precision defined as follows:

$$\text{Recall} = \frac{tp}{tp + fn} \qquad\qquad \text{Precision} = \frac{tp}{tp + fp}, \tag{6}$$

where tp is the number of correctly recommended relevant items, fn is the number of wrongly recommended relevant items, and fp is the number of wrongly recommended non-relevant items. We report Recall@N and Precision@N which is the Recall and Precision in a ranked list just considering the first N items.

Baseline method. As a baseline, we use the classic *Bag of Words* approach and represent the free text descriptions of the clothing items as TF vectors as opposed to CF in our proposed approach.

4.3 Results

We report the results in Fig. 4. For the evaluation metrics, we focused on N = 5, since our system will display 5 recycling recommendations to the user. From the figure, we note that our proposed approach using the CF weighting scheme outperforms the baseline in both categories.

5 Related Work

Smart closets. Previous works [8,9] have proposed similar architectures with RFID-enabled clothing items that is leveraged to generate outfit recommendations. Some other earlier works [10,11] are also leveraging RFID technology—as shown in this paper, but the applications in these works are limited to inventory overview. In this paper, we mainly focus on leveraging our proposed architecture to recommend textile recycling suggestions.

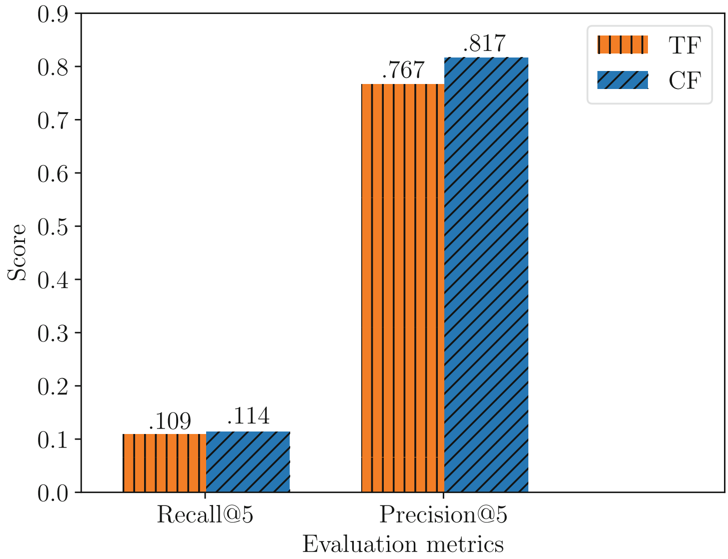


Fig. 4. Experiment results of Recall@5 and Precision@5 for the baseline (TF) and our proposed approach (CF). The reported results are the average of Recall@5 and Precision@5 for each user evaluated in the experiment.

Recommendation approach. In recent years, LOD in recommender systems has been frequently researched and various applications have been proposed [4,12]. The most common application is to calculate semantic similarity of items based on the item's relationships found in datasets published as LOD. Using concepts from a LOD Knowledge Base to model the user profile have shown promising results in past work, e.g., [13]. Many of these past works use `DBpedia`[4] as a LOD Knowledge Base and focus on recommendation in the traditional domains where large datasets are available, such as movies, music, and books.

Our recommender system addresses a relatively unexplored domain and exploits a LOD Knowledge Base lacking previous research. To the best of our knowledge, recommending items that are no longer of interest to the user is a quite recent idea. Moreover, with the proposed approach built into the architecture of the Internet of Things wardrobe, this paper's proposed recommender system in the fashion domain is the first of its kind.

6 Discussion

CF vs. TF. In the following, we highlight the advantages our proposed CF approach has over the TF baseline. These highlights should give an idea of why the CF outperforms TF. A drawback of *Bag of Words* is that it considers all

[4] http://wiki.dbpedia.org/.

words in the text descriptions as equally important. To use the item in Fig. 3 as an example, it represents the item as an item containing the word 'miu' two times, when—in fact—'Miu Miu' is the brand of the item and is of vital importance to the representation of the item. For clothing items with brand names of common terms, such as 'Jean Shop', the *Bag of Words* approach is in deep trouble. In advantage, the *Bag of Concepts* approach is able to extract the 'Miu Miu' entity (`Q552230`) and able to capture the semantic context of the item. Moreover, the *Bag of Words* approach will describe the item using words such as 'clean', which does not characterize the item in any way.

Textile recycling recommendations. By recycling textile waste, benefits can be seen in the environmental, the economical, and the social sustainability. With today's recommender systems already making a huge impact on people's pre-purchase behaviour, it is evident that they also can affect people's post-purchase behaviour. A system architecture and a recommender system as described in this paper, is a good candidate to affect people's post-purchase behaviour for the social good, by recommending clothes that the system's users can remove from their wardrobes.

7 Conclusion and Future Work

In this paper, we describe an Internet of Things wardrobe enabled with a proposed semantic content-based recommendation approach to recommend clothing items for recycling. We describe a set of context signals obtained from the wardrobe's architecture and how they affect the recommended list displayed to the user. Evaluation of our approach shows that our approach outperforms a baseline in terms of accuracy. Moreover, previous research has shown that LOD can increase recommender system's transparency and increase the user's trust in the system by computing convincing explanations to the recommendations [15]. The proposed approach facilitates opportunities for this and is planned for later research. Hence, the approach poses as a promising fit for the system.

Future work will also be devoted to improving the recommendation approach with a *Concept Frequency - Inverse Document Frequency* weighting scheme, and to develop further steps in system for the users to act on the recommendations.

References

1. Chavan, R.B.: Environmental sustainability through textile recycling. J. Text. Sci. Eng. 1–6 (2014). https://www.omicsonline.org/special-issues/textile-recycling-its-applications.html
2. Klepp, I.G., Laitala, K.: Clothing consumption in Norway. Technical report, Consumption Research Norway (SIFO), in Norwegian (2016)
3. Lops, P., Gemmis, M.D., Semeraro, G.: Content-based recommender systems: state of the art and trends. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P. (eds.) Recommender Systems Handbook, pp. 73–105. Springer, Boston (2011). https://doi.org/10.1007/978-0-387-85820-3_3

4. Noia, T.D., Mirizzi, R., Ostuni, V.C., Romito, D., Zanker, M.: Linked open data to support content-based recommender systems. In: Proceedings of the 8th International Conference on Semantic Systems, I-SEMANTICS 2012, pp. 1–8. ACM, New York (2012)
5. Bizer, C., Heath, T., Berners-Lee, T.: Linked data-the story so far. Int. J. Semantic Web Inf. Syst. (IJSWIS) **5**(3), 1–22 (2009)
6. Kolstad, A., Özgöbek, Ö., Gulla, J.A., Litlehamar, S.: Connected closet - a semantically enriched mobile recommender system for smart closets. In: Proceedings of the 13th International Conference on Web Information Systems and Technologies, WEBIST 2017, Porto, Portugal, pp. 298–305 (2017)
7. Kolstad, A., Özgöbek, Ö., Gulla, J.A., Litlehamar, S.: Rethinking conventional collaborative filtering for recommending daily fashion outfits. In: Proceedings of RecSysKTL Workshop @ ACM RecSys 17, RecSysKTL 2017, Como, Italy, pp. 22–29 (2017)
8. Goh, K.N., Chen, Y.Y., Lin, E.S.: Developing a smart wardrobe system. In: 2011 IEEE Consumer Communications and Networking Conference (CCNC), pp. 303–307, January 2011
9. Yu-Chu, L., Kawakita, Y., Suzuki, E., Ichikawa, H.: Personalized clothing-recommendation system based on a modified bayesian network. In: 2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet, pp. 414–417, July 2012
10. Toney, A.P., Thomas, B.H., Marais, W.: Managing smart garments. In: 2006 10th IEEE International Symposium on Wearable Computers, pp. 91–94, October 2006
11. Ling, S., Indrawan, M., Loke, S.W.: RFID-based user profiling of fashion preferences: blueprint for a smart wardrobe. Int. J. Internet Protoc. Technol. **2**(3/4), 153–164 (2007)
12. Figueroa, C., Vagliano, I., Rocha, O.R., Morisio, M.: A systematic literature review of linked data-based recommender systems. Concurr. Comput.: Pract. Exp. **27**(17), 4659–4684 (2015)
13. Piao, G., Breslin, J.G.: Exploring dynamics and semantics of user interests for user modeling on Twitter for link recommendations. In: Proceedings of the 12th International Conference on Semantic Systems, SEMANTiCS 2016, pp. 81–88. ACM, New York (2016)
14. Sahlgren, M., Cöster, R.: Using bag-of-concepts to improve the performance of support vector machines in text categorization. In: Proceedings of the 20th International Conference on Computational Linguistics, COLING 2004, Stroudsburg, PA, USA. Association for Computational Linguistics (2004)
15. Musto, C., Narducci, F., Lops, P., Gemmis, M.D., Semeraro, G.: ExpLOD: a framework for explaining recommendations based on the linked open data cloud. In: Proceedings of the 10th ACM Conference on Recommender Systems, RecSys 2016, pp. 151–154. ACM, New York (2016)

Games, Assessment and Rehabilitation: When Serious Games Support Cognitive Development in Children with Cerebral Visual Impairment

Matteo Ciman[4], Ombretta Gaggi[1,3](✉), Teresa Maria Sgaramella[2,3], Laura Nota[2,3], and Margherita Bortoluzzi[3]

[1] Department of Mathematics, Università degli Studi di Padova, Padua, Italy
[2] Department of Philosophy, Sociology, Education and Applied Psychology, Università degli Studi di Padova, Padua, Italy
[3] Centro di Ateneo Disabilità Riabilitazione e Integrazione, Università degli Studi di Padova, Padua, Italy
{ombretta.gaggi, teresamaria.sgaramella, laura.nota}@unipd.it, margherita.bortoluzzi@hotmail.it
[4] University of Geneva, Geneva, Switzerland
matteo.ciman@gmail.com

Abstract. In this paper, we present a study on the use of serious games in the assessment and rehabilitation of children with Cerebral Visual Impairment (CVI). Moving objects support vision and processing the information conveyed in children with CVI; they also find it easier to deal with more simple images. Our serious games help them keeping focused on the exercise by using touch interface, the game paradigm and cartoon characters. The paper reports lesson learned from data collected in a user study to highlight the high potential of using these games also in the rehabilitation process, which brings us to develop the games also in a mobile platform to allow children train the skill at home, i.e., more intensively and in a familiar environment.

Keywords: Serious games based assessment and rehabilitation
CVI

1 Introduction

Cerebral Visual Impairment (CVI), also known as Cortical Visual Impairment, is a disability that entails a visual deficit, due to a brain damage (Roman-Lantzy 2007). People with CVI are not able to see an object if this is not moving. Moreover, they have a reduced ocular field and ocular delay, find difficult to understand complex images and they are not able to see and touch an object at the same time. Child affected by this disability can experience different levels of these difficulties.

Many studies have shown that Serious Games can help children affected by disabilities since they help to obtain their attention and to protract the training session

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 61–70, 2018.
https://doi.org/10.1007/978-3-319-76111-4_7

since they have fun (Gaggi and Ciman 2016; Gaggi et al. 2017). In Ciman et al. (2013) and Gaggi et al. (2016) we described two serious games designed to train children affected by CVI. The system allows tracking both eyes movements, thanks to an integrated eye-tracker module, and users' touch interaction. *CathMe!* is a Serious Game which asks children to follow a cartoon character chosen according to child's preferences by both looking to it and touching it. *HelpMe!* asks users to help Santa Claus's assistant classify presents according to their category. Each level of the game has a different set of target semantic category, e.g., "musical instruments", "clothes for dolls", "cars" and so on; each of them is composed by a set of different object images belonging to the specific semantic category. Children are required to recognize among intruders objects belonging to the target category announced by the speaker.

We initially tested the system with 28 children from kindergarten, aged from 3 to 6 years old, who did not experience disability. This initial testing confirmed the acceptability of the system, and the confidence of the children while playing with it. Information about performance in children with normal sight level were also reported in Gaggi et al. (2016) and were used as base line in this paper when analyzing performance in children with CVI. A second test was conducted by asking 19 children affected by CVI to play with our games. In this paper, we present an example of a specific pattern emerged with one participant to highlight the high potential use of these games with the results obtained with a four years old child.

The paper is organized as follows: Sect. 2 describes performance and, more specifically in Sect. 2.1, results obtained by one of the study participants. In Sect. 3 we describe the implementation of the games in the mobile platform to allow children to train at home. Finally, Sect. 4 presents conclusions and directions for future research studies.

2 Strengths and Rehabilitation Perspectives of Serious Games Assessment Tools: The Example of a 4 Years Old Child with Marked CVI

As already mentioned, the test phase involved 19 children affected by CVI. Children involved were aged between 4 and 9 years old, 13 of them were females, 5 were males. Unfortunately, not all the data collected during the training sessions are significant to present for several different reasons: first, some children with CVI are unable to move their hands in a coordinated way, so they are not able to keep the finger touching the screen along the activity proposed; second, sometimes the eye-tracker was not able to collect data due to the child difficulty in maintaining the head in the correct position, or because of too many head's movements; finally, the variability of children's age with respect to data available from children without CVI limit the comparison with peers without CVI. For these reasons, in order to highlight the potential of this type of assessment, we decided to provide the analysis of a single profile as an example, presenting changes in her performance in both the activities proposed and comparing her responses with data currently available from age-matched peers. Our analysis highlights a series of information emerging from the performance of Hope, a 4 years old girl with a marked CVI and residual vision who was able to play for the full session.

Hope has been proposed our serious games, which consist in different interactive situations through which the typical development and assessment of visual abilities in children is conducted. The goal of the proposed activities was twofold, i.e., to assess the level of processing of the visual information available to Hope, and to analyze the learning potential and retention of visual information.

In particular the proposed tasks were:

A. *Where is Nemo going?* Hope practiced with movements in the visual field. She was asked to look to the goldfish Nemo, which will appear on the screen. This game was used to calibrate the system.
B. *CatchMe! Where is Peppa/George going?* (Fig. 1, left) Hope have to follow a cartoon character chosen according to participant gender, which was moving on the screen. The task goal is to orient her attention.
C. *HelpMe! The Santa Claus's assistant* (Fig. 1, right). She was asked to help Santa Claus in a long-lasting task, i.e., to prepare in advance the sack of Christmas presents. Pictures of objects belonging to three different target semantic categories (animals, vehicles and clothing) are used together with objects belonging to other categories (intruders). All pictures take into account the age of acquisition and the organization of the children semantic lexicon at her age. Hope was then required to orient the attention and focus on the object, to discriminate between target images (i.e., images that belong to the target category) and intruders, and make a cognitive decision putting the target object into Santa Claus' sack, and throwing out of the screen the other ones. If she failed to discriminate one of the images of a particular set of pictures, she was asked to repeat the actions. After three failures in the same level, the game moves to the next sequence.

Fig. 1. *CatchMe!* (left) and *HelpMe!* (right) games.

2.1 Results

Hope performance is described with respect to changes occurring along the execution and age matched peers in order to highlight strengths and vulnerabilities. While game

Were is Nemo going? was used to develop familiarity with the situation and calibrate the system, eye movements, accuracy of the answer and reaction times were collected and analyzed respectively for games *CatchMe!* and *HelpMe!*.

CatchMe! Visuo-motor coordination game following a visual stimulus moving on the screen
The goals set with this game were twofold. First of all it aimed at describing the basic level of the target skill, that is visuo-motor coordination; secondly at providing information on the learning potential of the participant in order to make more effective subsequent rehabilitation decisions. To accomplish the second goal the game was proposed a second time and some manipulations were introduced (Sgaramella 2016).

Figure 2 represents the Euclidean distance between the center of the image presented on the screen and, respectively, the position of Hope's eye (Gaze) and the position of the finger on the screen (Touch). The green solid line represents the position of the Target picture on the screen, which however changes over the task. The picture provides a temporal description of Hope performance occurring across the activity and highlights the direction of changes occurring for an interval of time of 30 s. Along the time interval described in the graph, the Touch index results in an almost stable line with a more marked reduction in the last portion of the considered time interval. At the same time, the picture shows a decrease in the distance between the point of the actual touch and the position of the picture on the screen, with the gap passing from more than 1000 pixels to an interval comprised between 500 and 250 pixels. Hope's finger is then closer and closer to the picture[1], thus suggesting a systematic trend toward a more accurate performance and a learning process activated in a visuo-motor task.

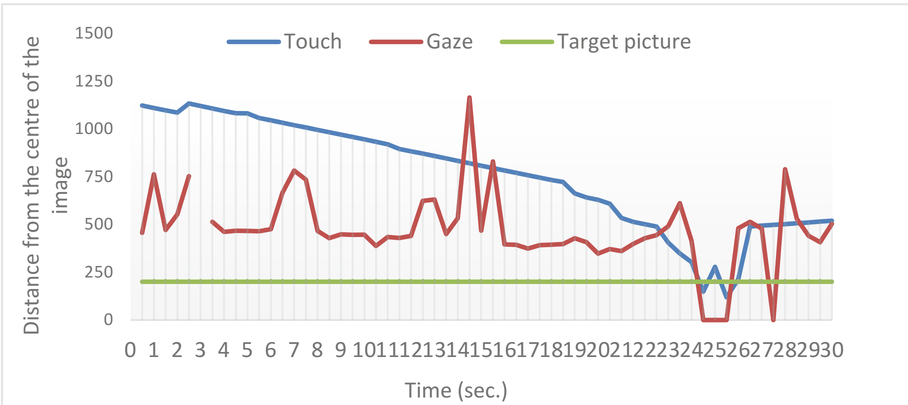


Fig. 2. Delta for Touch and Gaze with respect to the target picture in a portion of task execution

[1] Consider that the distance is computed from the center of the picture, but the picture itself has a dimension, which depends on the user's setting, and in this case is 200×400 pixels.

Sustaining learning with repeated task execution and a higher complexity
Hope was then required to execute the task for a second time. In this case, starting from a higher performance gained in the first execution, she was presented the target item in different portion of the screen, hence requiring her to orient her visual attention to different portions of the visual field.

In Fig. 3, which is focused on the gaze analysis, the distance between the Gaze line and the solid line representing the Target image was rapidly and markedly decreasing along the time interval described. The finger pointing was consistently close to the actual position of the object moving on the screen. Therefore, in the second trial, once she refreshed the task, her performance improved in accuracy and persisted at very accurate level thus suggesting that continuing with the training the performance could be more accurate and stable displaying her learning potential in focusing attention and in coordinating the hand movements in order to accomplish the task.

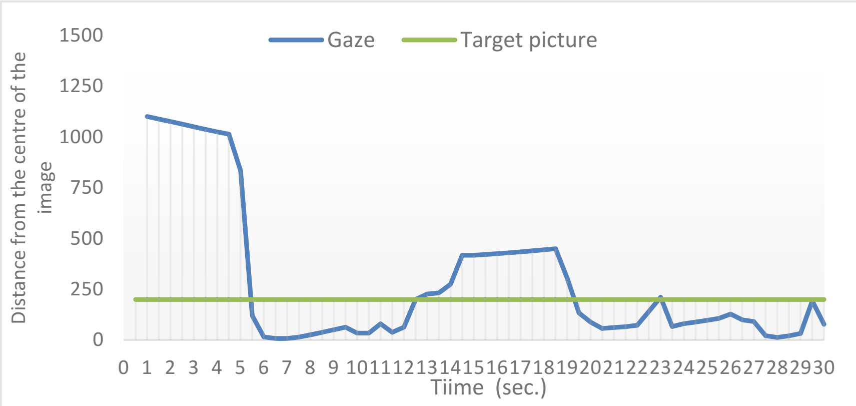


Fig. 3. Delta for Gaze with respect to the target picture in a portion of the repeated task execution

***HelpMe!* A controlled access to complex visual information in a visual decision making task**
Given the relevance of the processes investigated in the cognitive development of children and the relevance for school inclusion, cognitive decision making relying on more complex visual information is introduced. Both accuracy of visual attention performance and decisions made are then analyzed.

Accuracy in visual semantic decision making
In this game, Hope was required to decide on the sematic category to which the object represented on the screen belonged to. Her performance was 77% of correct answers in the first presentation of the items. However, she was capable of reaching full correct performance in the second execution of the task.

Figure 4 shows the details about her performance as compared to age matched peers and described in terms of standard deviation from mean level obtained with 28 children in a previous study (Gaggi et al. 2016). The decision time refers to the time needed to take a Yes/No decision about the semantic category of the object appearing on the screen, e.g., the time that passes between the image arrives on the screen and the first interaction of the child with that image to move to the right place; the completion time refers instead to the time needed to move the object from the position where it has appeared to the Santa's sack, or to put it in a corner.

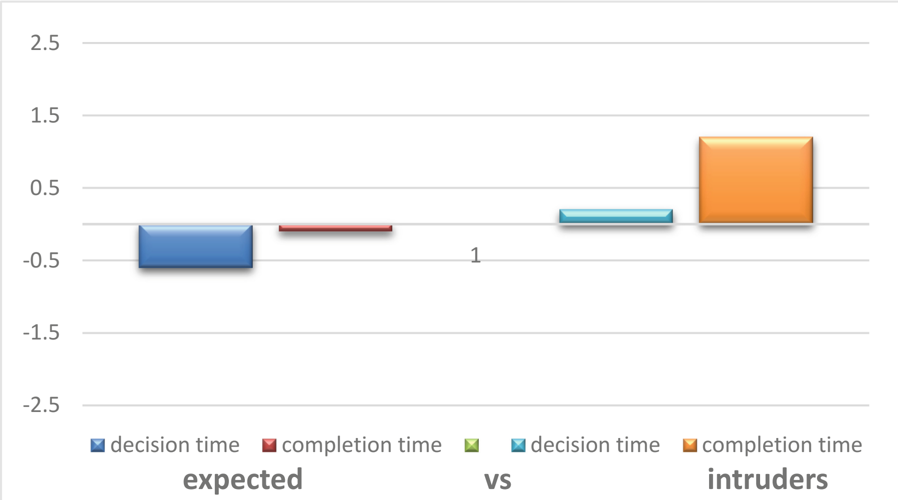


Fig. 4. Hope reaction times expressed as standard deviation from age matched peers in the decision and completion task for both expected items and semantic intruders

The game highlights Hope ability in selecting the appropriate object, maintaining the gaze to an attentional level adequate to keep track of the movement until she reaches the appropriate destination. The time required both to decide and to complete the task which was expressed in milliseconds, is within the normal age range insofar it falls within ±1.5 standard deviation. This applies both when the showed picture belongs to the category announced by the speaker and when it is an intruder. Finally, deciding that an object is an intruder, that it does not belong to the semantic category, is a more complex process and thus requires a longer time.

Category specific decision making

Objects proposed to Hope belonged to different semantic categories. Her specific skill in making decision about objects from different categories, also characterized by different levels of internal visual structure (e.g., an animal or a car) was then analyzed. Additionally, her answers to both expected items and intruders, requiring a specific decisional ability, were separately analyzed. Figure 5 shows Hope performance as compared to age matched peers.

As shown in the figure, items from different categories, with different internal structure, showed a similar level of difficulty for Hope, with the performance in decision making falling within normal age range, either for expected and intruders, that is with unexpected items.

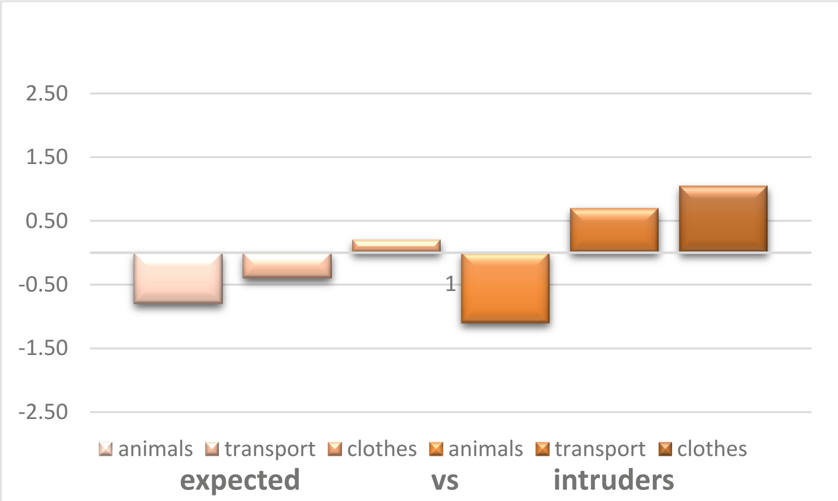


Fig. 5. Hope reaction times expressed as standard deviation from times obtained by age matched peers in the decision making task for both expected items and semantic intruders

Completion time for specific semantic categories
The time needed to complete the task, for different semantic categories was then analyzed.

As shown in Fig. 6, Hope's performance was within normal age range with expected items, i.e., objects congruent with the category announced by the speaker. With intruders of different categories, the time required to complete the task was significantly longer thus suggesting a longer execution time but only when it deals with a structurally more complex category such as means of transportation, although items were characterized by the same familiarity level and age of acquisition as a semantic category.

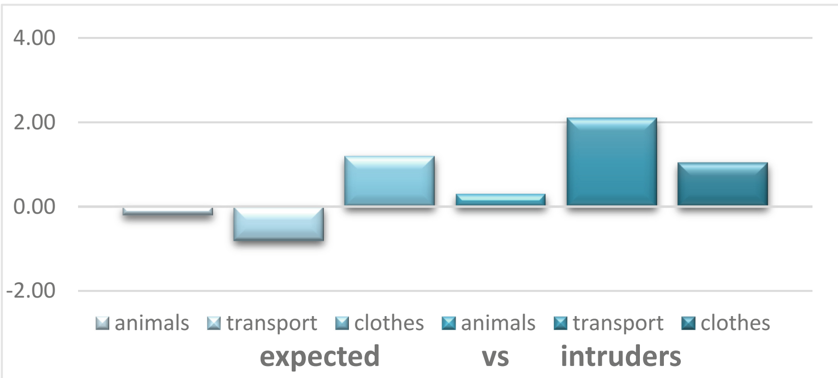


Fig. 6. Hope reaction times expressed as standard deviation from times obtained by age matched peers in the completion task for both expected items and semantic intruders

Summary of results
As far as Hope is concerned, data collected with these games highlighted her visual attention and visuo-coordination skills but also visual object recognition and visuo-semantic decision making. Additionally, these games have been useful in describing the improvement in performance during the task, as well as her ability to follow a moving object keeping her gaze coordinated with the finger till the object reaches the final appropriate destination. Indexes used were then useful in tracking learning development in terms of both accuracy and speed of processing, which is relevant in terms of rehabilitation choices and for future perspectives.

From a methodological level, the assessment described supports the adequacy of the tools, at least when working with 4 to 6 years old children with CVI and residual vision, either in the assessment and in supporting the development of decision making skills. It provides guidelines for subsequent applications of the tools, that is: proposing repeated execution of the tasks, looking at both accuracy and reaction times, manipulating complexity either in visual scanning of the screen and in complexity of the visual structure of the pictures used.

3 Going Mobile for Ecological Intensive Rehabilitation/Practice

As we already presented in Sect. 2, results from the assessment of specific visual skills in a child with CVI, when tested with our system, were useful both in assessing basic skills and in analyzing learning potential and persistence of learned tasks related to the ability of recognizing and following object moving on the screen and, more demanding, on taking cognitive decisions on the incoming materials.

Considering that the described results are observed in a test that lasted for a very short interval of time, about 15 min, the possibility of using the game for longer time intervals becomes very interesting and more appealing. A further step is then implementing a solution, which might allow children and their parents to have the opportunity of practicing these activities in a more extended, and at the same time intensive way. The idea is to give the children the possibility of using the game at home, thus giving the possibility of practicing the more they can, whenever they want, with the support of their parents, and in a more comfortable and familiar environment. Since the games are proved to be engaging, either the duration of the training session and the complexity can be set and modified when needed to a more complex and demanding level of analysis thus providing a personalized learning.

For this reason, we implement a mobile version of our games *CatchMe!* and *HelpMe!*, to make them available to children to play at home, maybe following a program provided by their doctor.

One of the problems which is necessary to address when developing applications for mobile users is related to market fragmentation. Despite the low number of children affected by CVI, supporting only a particular device/operating system could strongly reduce the number of potential users, hence reducing the possible benefits of our application. For this reason, the best approach for developing is to use a cross-platform framework (Ciman and Gaggi 2017), that let developers build only one application using the framework specific language, and deploy this application to multiple mobile operating systems, e.g., iOS, Android, Windows Phone, etc. Since the system was initially developed as a web application, the best candidate solution was Apache Cordova framework[2], which allows using a webkit engine to wrap a web application into a mobile hybrid application that can be installed on mobile platforms.

Since we cannot assume that all children have a smartphone or a tablet and since they surely do not have an eye-tracker at home, when they play at home we can collect data about touch but not about gaze.

The mobile application, after requesting for username and password defined with the doctor when the child was added to the system, let the child decide with which game he/she wants to play, i.e., *CatchMe!* or *HelpMe!* Then, the application downloads the game settings for the child, that could be adapted by the doctor after having analyzed the performances of the child. Then while the child is playing the game, the application records information about the touches and the interaction of the child with the game, and stores this data locally in the memory of the application. Only at the end of the game session, or when an internet connection is available, the application will ask to upload the data on the server. The uploaded data contain information about the game settings used during the game played by the child, e.g., screen resolution, images size, etc., the interactions of the child with the screen and the game and, in case of the *HelpMe!* game, the number of correct and incorrect answers. For the *CatchMe!* game, once all the data upload has been completed, the server will analyze interactions and image's movements to provide an evaluation of the touch performances of the child.

Once uploaded, the doctor is able to analyze data and performance of each child for the specific exercise carried out; he/she can also modify settings and tasks to make them more suitable for further progresses of the child and/or taking into account the difficulties encountered by the child while exercising at home.

4 Conclusions

In this paper, we have analyzed the results obtained using our serious games for the assessment and rehabilitation of children affected by CVI. Combining both basic visual search task with more complex tasks requiring attention and cognitive processes, such as decision making, allows a 'whole child' evaluation and helps identifying all areas of strengths and difficulties, a direction advocated by neuropediatrics as crucial for fostering the development of children with CVI.

By supporting the assessment of preserved skills, the system supports also a personalized learning for children with CVI, that is providing a condition in which the pace of learning, the level of difficulty and the instructional approach are optimized for the needs of each learner. In addition, the activities implemented in the system are meaningful and familiar to learners thus driving their attention. Identifying ocular, oculomotor, perceptual and/or visuo-cognitive preserved skills is, in fact, essential for

[2] https://phonegap.com/.

CVI children in order to apply specific habilitation programs and increase their potentials for inclusion and participation.

Despite the difficulty in collecting data that can be compared with normally developing children, the analysis of data collected with a four year old child has shown an improvement in the performance even with a very short training session (only 15 min). Although at the moment supported by a single user, the system is promising for implementing a personalized learning. Further experimentation is also needed for instance with gradual inclusion of additional features. The mobile version of our serious game, which we implemented, can be effective in allowing children to train themselves at home, for longer interval of time. Moreover, the system implemented provides a mechanism for recording learning achievements and communication between the family and the professionals. The pace and learning development curve can provide the basic information for identifying strategies for everyday life both at home with family members and for teachers in the educational settings.

References

Ciman, M., Gaggi, O.: An empirical analysis of energy consumption of cross-platform frameworks for mobile development. Pervas. Mob. Comput. **39**, 214–230 (2017)

Ciman, M., Gaggi, O., Nota, L., Pinello, L., Riparelli, N., Sgaramella, M.T.: HelpMe!: a serious game for rehabilitation of children affected by CVI. IN: Proceedings of International Conference on Web Systems and Technologies (WEBIST 2013), pp. 257–262 (2013)

Gaggi, O., Ciman, M.: The use of games to help children eyes testing. Multimed. Tools Appl. **75** (6), 3453–3478 (2016). Springer, Netherlands

Gaggi, O., Palazzi, C., Galiazzo, G., Franceschini, S., Ruffino, M., Gori, S., Facoetti, A.: Serious games for early identification of developmental dyslexia. ACM Comput. Entertain. **15**, 4 (2017)

Gaggi, O., Sgaramella, T.M., Nota, L., Bortoluzzi, M., Santilli, S.: Serious games system for the analysis and the development of visual skills in children with CVI. In: Gaggi, O., Manzoni, P., Palazzi, C., Bujari, A., Marquez-Barja, J. (eds.) GOODTECHS 2016. LNICS(SITE), vol. 195, pp. 155–165. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-61949-1_17

Roman-Lantzy, C.: Cortical Visual Impairment: An Approach to Assessment and Intervention. American Foundation for the Blind, New York (2007)

Sgaramella, T.M.: Ri-abilitazione: linee guida e prospettive [Re-habilitation: guidelines and perspectives]. In: Soresi, S. (ed.) Psicologia delle disabilità e dell'Inclusione [Psychology of Disabilities and Inclusion]. Bologna, Il Mulino (2016)

How Blind People Can Manage a Remote Control System: A Case Study

Marina Buzzi[1], Francesco Gennai[2], and Barbara Leporini[2(✉)]

[1] IIT-CNR, via G. Moruzzi, 1, 56124 Pisa, Italy
marina.buzzi@iit.cnr.it
[2] ISTI-CNR, via G. Moruzzi, 1, 56124 Pisa, Italy
{francesco.gennai,barbara.leporini}@isti.cnr.it

Abstract. Remote Control Systems (RCSs) are increasingly being installed in homes and offices. Technology evolves very rapidly and sensors and devices are becoming smaller, smarter and more powerful. Mobile and Web apps are commonly used to remotely configure and control devices. Home control is especially valuable for blind people, since they can benefit from technology to control and turn on/off devices autonomously. Remote control can offer meaningful support, especially when devices are not directly accessible (e.g., thermostat to manage heating temperature). Therefore, if RCS interfaces are not accessible via screen reader, blind users may miss a great opportunity to achieve greater autonomy at home. This paper investigates the accessibility of the web user interfaces offered by RCSs for blind people. To do this, the Fibaro, a popular Remote Control System, was tested as a case study by analyzing the interaction via screen reader. Results indicate that accessibility and especially usability need to be improved to make interaction easier and more satisfying for blind people. To this aim, some suggestions are offered to aid developers in designing more accessible RCS user interfaces.

Keywords: Home Automation · Accessibility · Blind users

1 Introduction

The last decade has been characterized by great advances in Home Automation and Remote Control Systems (RCSs), increasing the power of devices and sensors while decreasing their size and costs. These smart systems are a great opportunity for the autonomy and safety for elderly and disabled persons, enabling one to control lights, shutters, doors, alarm systems, smart TV, cameras, sensors, devices, and so on. Visually impaired people, and especially the blind, can increase their autonomy enormously, thanks to technology that supports controlling and handling devices via smartphone, tablet or computer. Many devices that are not directly accessible can today be handled via mobile or Web-based applications also by a blind person, making possible various activities, such as programming a washing-machine or setting the household heating temperature, if the user interfaces (UIs) and functionalities are appropriately designed. Briefly, Home Automation Systems (HAS) are today easily controlled via smartphones, tablets or computer with customizable apps. However, to

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 71–81, 2018.
https://doi.org/10.1007/978-3-319-76111-4_8

offer benefits and enable everyone to take advantage of the full potential of these systems, the Web-based or mobile apps for remote control should be accessible and usable also via assistive technologies, such as a screen reader for blind persons.

In this paper we investigate the accessibility and usability of a Web application used to remotely control components of a Home Automation System. A popular HAS was selected as a case study. Of several commercial systems, we selected the Fibaro (https://www.fibaro.com), relying on Z-Wave protocols, which offers a system for managing a building accessible via browser through a graphic user interface (GUI) or via the native mobile app. It includes an extensive library of preinstalled plugins for third-party smart TVs, cameras, thermostats, alarm panels, media players and control systems. An easy-to-use visual scene builder can be used to configure the system. The Geo-Location feature enables the creation of GPS triggers for all users anywhere in the world. Different degrees of personalization can be achieved by advanced users who are able to write scripts with Lua, a powerful and efficient scripting language (http://www.lua.org).

Our goal is to investigate (1) the main components of an RCS suitable for simplifying interaction for blind people, and (2) the problems a blind person might experience when interacting via screen reader with a HAS. Regarding the main components of a remote control system, blind persons have different or additional needs compared to sighted people. Since a blind user cannot have an overview of the interface, a sequential reading via a screen reader is necessary. This can require considerable effort. Therefore, appropriate UI design can improve interaction via screen reader as well as via keyboard. For example, a functionality designed for making a list of the only rooms with open shutters could dramatically simplify and shorten the interaction via screen reader when the user wants to close the opened shutters. Such a functionality might not be available in a common HAS since a sighted person spends a very short time exploring the UI at a glance to check the rooms that still have open shutters. Other similar functionalities could be very useful to greatly improve the usability of the interaction for the blind. Concerning interface accessibility, to ensure simple interaction for anyone, regardless of his/her disability, it is important to first implement the Web Content Accessibility Guidelines (WCAGs) including structured content, accessibility tags, and the possibility of managing any actions via keyboard, in order to avoid obstacles or difficulties experienced by a blind person. For example, links and buttons should be context-independent and auto-explicative (e.g., “Turn light on” rather than simply “Turn on”). In the following, we analyze the Fibaro web user interface to investigate the aforementioned aspects. The paper is organized as follows: Sect. 2 reports related work, Sect. 3 defines the methodology and introduces the study. In Sect. 4 results of the inspection via screen reader of the Fibaro Web interface are discussed and some suggestions are proposed for developers of Home RCSs. Conclusions and future work end the paper.

2 Related Work

Nowadays, low-cost Home automation and security systems can be implemented using Android (the popular Operating System for mobile devices) and Arduino (Arduino is an open-source electronics platform based on easy-to-use hardware and software, www.arduino.cc), to control home appliances and provide security via phone or tablet [4]. However, since it may be difficult for blind users to deal with these open systems and considering that the market offers low-cost built-in solutions, such as Fibaro, relying on open languages, we selected this system in our study.

Smart Houses take advantage of the Internet to make home automation smarter and easier to use anytime, anywhere. The inclusion of health systems to assist elderly and disabled people furthers technology's progress toward the Health Smart Home (HSH) [6, 10]. As suggested more than a decade ago by Stefanov et al., "Some innovations in smart houses designed for people with disabilities (such as advanced health monitoring systems) may be transferred to non-disabled people. Wearable monitoring systems will become important tools for prolonging life expectancy" [10]. Actually, simplifying eye-free interaction with home RCS systems could benefit anyone in specific conditions (e.g., no light or temporary disability).

With this in mind, Santos et al. introduced B-Live, a Home Automation System specifically designed for disabled and elderly people, to enable a conventional home to become an HSH. The system has been tested with motor impaired users (tetraplegic, paraplegic and in wheelchairs) at their homes [9], enabling users to simply carry out operations such as turn on/off the room lights and open/close room blinds and doors. Sandweg et al. designed a telephone-based interface (TBI) in the development process of a home automation system, collecting user requirements in focus groups and applying general guidelines for TBI design. The introduction of no-speech sounds may improve the user experience [8].

An accurate review on AAL is carried out by Al-Shaqi et al. Specifically, authors analyzed the last 15 years of research, discussing the current practices for developing AAL systems and identifying future research directions, commercial, technical and social challenges [1]. Portet et al. carried out an experiment for assessing the acceptability of a smart home equipped for vocal interaction (audio processing technology) with a sample of autonomous elderly people living alone in their home. Results showed that speech technology has a great potential to ease everyday life. However, the system encourages a lazy lifestyle, provoking a rapid degradation of health. Smart home technology design must promote a healthy way of living [7].

Domingo offers an extensive view of the potential of the IoT (Internet of Things) in different scenarios, including the smart home. Home devices include sensors and actuators embedded in goods, home appliances or furniture. Smart homes for people with disabilities provide accessible interfaces to manage the home devices for automation and control and include assistive devices to improve autonomy, to counter their social isolation and monitor their health [3]. As previously stated, many studies investigate AAL systems. Several of them focus on solutions for the elderly, and those with disabilities in general. The needs of blind people have been thoroughly investigated and solutions have been designed to help them in everyday life (VizWiz Social) [2]. However, accessibility of interaction with RCS and HAS have still not been thoroughly investigated in all aspects (i.e., screen reader interaction).

3 The Study

The Fibaro system was configured to remotely control sensors and devices installed in the house of one of the authors. Some main functions have been selected in order to test the interaction via screen reader to investigate the following factors, valuable for a blind person. The user should be able to quickly and easily (1) check which devices/sensors are on/off; (2) turn the devices/sensors on/off; (3) get an overview of information about the home/room/device status. To this end, we selected some of the view modalities offered by the system: by (a) rooms, (b) devices, and (c) a single room (we selected the kitchen). Through each modality, the user should be able to obtain various data and have different opportunities to check and handle the home sensors' status. We focused on aspects to consider when simplifying interaction for a blind user. To interact with the user interfaces, the screen reader Jaws for Windows 18.0.2530 and Internet Explorer 11.00 was used. With regard to the screen reading interaction with the RCS, we specifically focused on the usability principles and not only on mere accessibility technical requirements. According to the usability and accessibility criteria proposed in [5], we selected some of them to evaluate the UI: Buttons and links labels, Page title, Page structure, Main content identification. The evaluation was conducted by the authors (one is blind) and two blind users with long experience in using the screen reader Jaws while surfing the Internet.

3.1 Case Study: The Fibaro System

Exploring the home page (Fig. 1), no particularly useful information can be accessed via screen reader. At the beginning of the page the user can obtain general information on the environment (like that reported in Table 1). Unfortunately, the information is not very clear; i.e., content like "Link Dashboard: 0/12 | 1/19 | 3/45 | L:0.00W | 14" results somewhat meaningless if not presented in a more user-friendly way. In addition, the screen reader user only encounters a list of rooms with no useful information associated. Therefore, the home page does not deliver helpful information. Besides, the graphical buttons in the main area of the interface are not detected as buttons but as links. Although this does not significantly affect the interaction, it is not semantically correct. Modern web pages rely on HTML and Javascript. To make images used as buttons to be perceivable via screen reader, a role should be associated using a specific ARIA tag (Accessible Rich Internet Applications) and the state of the object has to be managed programmatically [11]).

Case 1: Rooms View (Fig. 2). Table 1 shows how the screen reader Jaws interprets when one clicks on the link "rooms". The page is rich in elements. The sequential reading may take a long time. The user perceives a "long" list of the available rooms in the house. First, only a list of the rooms is available to allow the user to define a showing order to the rooms. Two links Up and Down are used for this purpose. This is not particularly useful for a blind user since when navigating via Tab key, too many Up and Down links are encountered; to know to what they refer to, the surrounding needs to be explored. Links should be more self-explanatory. Locating this function in a setting panel could make the system more usable. Moreover, on the page there is also

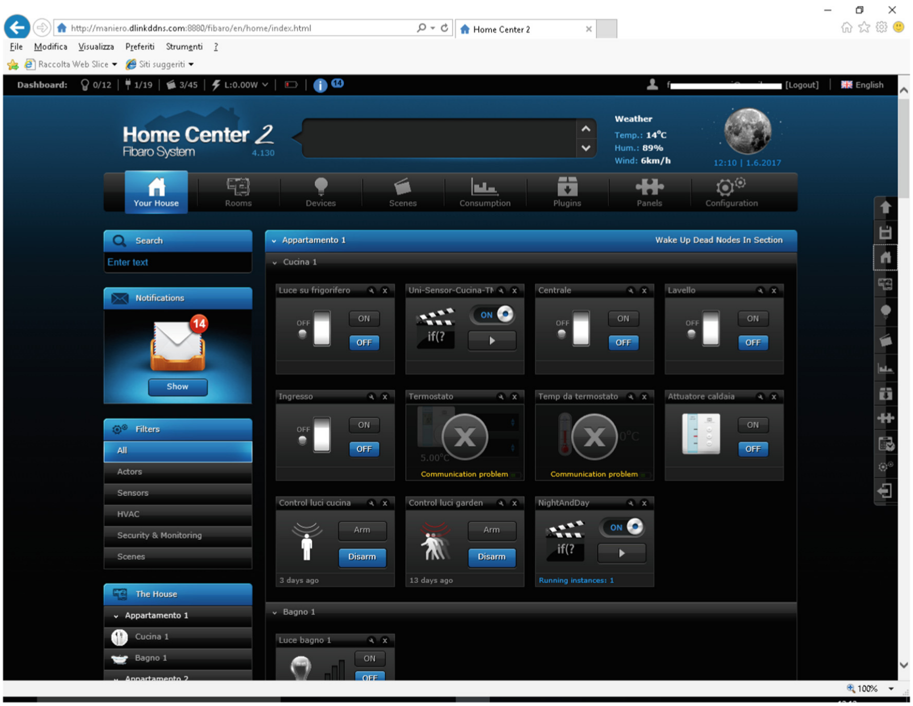


Fig. 1. Fibaro home page for a logged user

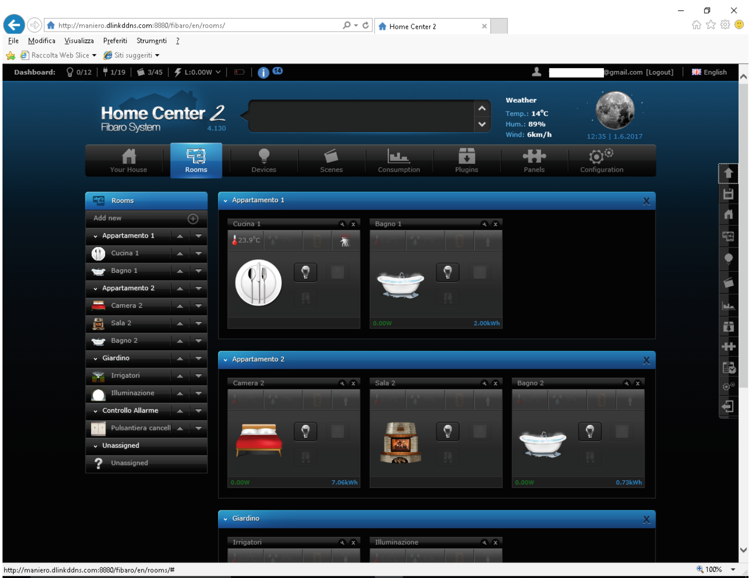


Fig. 2. Rooms view

Table 1. Rooms view read by Jaws: a fragment

Home Center 2
Link Up Save Your House Rooms Devices Scenes Consumption Plugins Panels Events Configuration Logout
list of 6 items
Link Dashboard: 0/ 12 1/ 19 3/ 45 L:0.00W 14
list end
name.surname@gmail.com [Link Logout] English
Link 4.130
Weather Temp.: 14°C Hum.: 89% Wind: 6km/h 12:38 | 1.6.2017
Link Your House Rooms Devices Scenes Consumption Plugins Panels Config
list of 5 items
Rooms
Link Add new
list of 4 items nesting level 1
Link Appartamento 1 Link down Link up
list of 2 items nesting level 2
Link Cucina 1 **Link** down **Link** up
Link Bagno 1 **Link** down **Link** up
...
Appartamento 1 rooms/#
Cucina 1 Edit Delete
23.9°C 0%
light/light0 roleta/roleta0
alarm/alarm_grey

information on the rooms. Each room is reported with some information: the name of the room followed by some numerical data. To reach each room the user needs to explore all the content sequentially. No mechanism for partitioning the content has been used (e.g., headings level, wai-aria regions, etc.). Moreover, the numerical data could be unclear for unskilled users. Analogous considerations about clarity are applicable to the similar Edit and Delete links.

Case 2: Device View. In Table 2 a portion of the page “Devices” content read by Jaws is shown. The available devices per room are listed (e.g., Bathroom 1 | list of 1 item | Bathroom 1 light | down up list end). No information about its status is available. Such a functionality is useful just to know the devices available in the room, but the user cannot do anything. The two links near each device (Up and Down) enable one to set up the device order. Therefore, when interacting via screen reader a blind person cannot obtain appropriate information or perform specific actions. A list of the available devices could certainly be useful to obtain an overview, but more useful information has to be captured via screen reader.

Figure 3 shows an overview of the sensors available on the house. All devices are listed and reported without a classification. The user needs to read all of them. If there are many sensors and devices, to check and manage a specific sensor or verify home parameters like the lighting status could require a lot of effort. Probably a sub-classification according to the type of sensors or devices could improve the usability of the remote automation system via screen reader.

Table 2. Devices page read by Jaws: a fragment

Link devices/#
Visibility
Link Show visible
Link Show hidden
Link Show all
Devices
Link Add or remove device
Link Cucina 1
list of 9 items
Link Luce su frigorifero **Link** down **Link** up
Link Centrale **Link** down **Link** up
Link Lavello **Link** down **Link** up
Link Ingresso **Link** down **Link** up
Link Termostato **Link** down **Link** up
…
Link Bagno 1
list of 1 items
Link Luce bagno 1 **Link** down **Link** up
list end

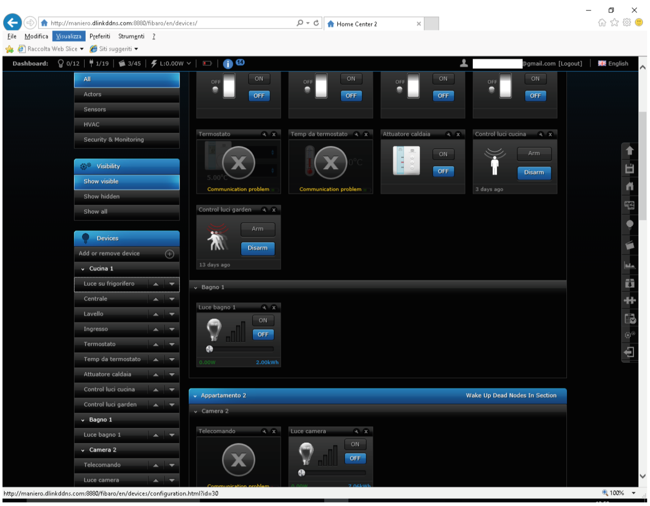


Fig. 3. Devices view

Case 3: A Single Room View. The information shown by the system when a specific room is selected is shown in Fig. 4a.

Even when choosing a single room, the interaction is not interesting due to too much content, no content partitioning, context-dependent links (Edit/Delete or On/Off), and unclear information. The editing of each room should be demanded in a setting panel. Moreover, although the two links On/Off associated to a lighting sensor allows the user to switch it on/off, no information on the current status (if the light is on or off) is

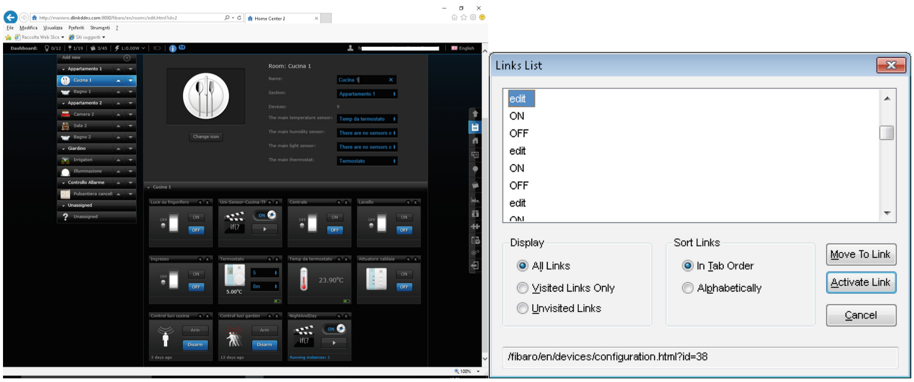


Fig. 4. (*a*) Single room (kitchen) view (*b*) list of links provided by Jaws

detectable via screen reader. Often blind users navigate interactive elements via Tab key or select links by a specific screen reader command (e.g., Insert+F7), moving via arrow keys. In Fibaro web interface labels are context-dependent; the element they are refer to is not clear (Fig. 4b) and thus it is necessary to explore the surroundings, increasing the interaction effort (Table 3).

Table 3. Single room (kitchen) view read by Jaws: a fragment

Cucina 1 **Link** rooms/edit
Link Change icon
Room: Cucina 1
Name:
Edit Cucina 1
Section: *Combobox* Appartamento 1
Devices: 9
The main temperature sensor: *Combobox* Temp da termostato
The main humidity sensor: *Combobox* There are no sensors of this type in the room.
…
Cucina 1
Luce su frigorifero **Link** *edit* **Link** delete
Link ON
Link OFF
Uni-Sensor-Cucina-TM **Link** *edit* **Link** delete
Centrale **Link** *edit* **Link** delete
Link ON **Link** OFF

4 Results and Suggestions

The following usability and accessibility issues emerged from the inspection analysis:

- Logical content partitioning: no content grouping/partitioning mechanism has been applied to the interface. This implies an "excessive sequence reading" and several difficulties to identify the main content of the current page. All the content is shown sequentially with no separation tags or main content identification.

- Meaningful page title: the page title (i.e., <title> tag) delivers useful information to screen reading users about the page content. The screen reader immediately reads the page title: if it is significant, the user can quickly get an idea about the page's main content. All the Fibaro explored pages have the same title.
- Labels: several links with the same content-depending labels, like "Edit", "Delete", "On", or "Off", are used. When the user navigates via Tab key, the links are not clear at all; the user needs to explore the nearby content to understand what those links are referring to. More significant and self-explanatory labels could help the user understand the link goal without exploring the surrounding. This makes the interaction more comfortable and effective.

These drawbacks negatively affect user interaction in satisfactorily handling the main functionalities such as:

- Check which devices/sensors are on/off: no dedicated button or filter to obtain a list of only the devices "on" is available. Also from the list of all the devices, the status information is not available since it is not provided in an accessible way.
- Turn on/off the devices/sensors: there is no function allowing one to set a specific category of devices (e.g., lighting). As a consequence, the user has to list all devices and one by one click "Off" (or "On"). This can require a lot of effort.
- Get an overview of information about the home/room/device status. Information on a given element can be difficult to obtain or can be incomplete. The pages are rich in information and focalizing on the desired topic may require a lot of effort or be difficult or impossible.

For blind people, home remote automation can be a valuable way to check and manage daily activities autonomously. For this reason, the interface should not only be accessible but also, and especially, usable. How the information on the devices as well as on sensors is presented may be crucial to facilitating the blind user's interaction.

Some access keys might facilitate the interaction, for example to activate the main push buttons, to obtain the device list, etc. In general, the interface is content overloaded, requiring a lot of effort during screen reader interaction. The configuration features should be grouped in a setting panel to lighten the interface.

Summarizing, using the standard interfaces a screen reader user is unable to obtain information about the status of devices and sensors, and to easily perform any action on them. Based on the results, we can offer some suggestions to keep in mind when designing the interface and configuration panel of a home remote system:

S1. Make the UI more readable and easy to interact with, i.e., by partitioning the content into several meaningful areas (WAI-Aria attributes and roles).
S2. Create self-explanatory links and buttons (labels), so that it is clear what they refer to. Use links "Kitchen light on" (or "Kitchen light off") rather than "edit", "on", or "off" as described in Example 3.
S3. Make different views available, such as by devices and rooms. For each view modality, the status of sensors and devices, as well as all types of contextual information should be presented in a clear and accessible manner.

S4. Provide functionalities to make some actions simpler, such as a button to set the status of all devices in the same typology (e.g., turn on all the lighting).
S5. Arrange all type of settings in a specific panel so that the user knows exactly when the content, links and buttons refer to a configuration.

5 Conclusions

The Results of this study pointed out some usability issues experienced by blind people when interacting via screen reader with the user interfaces of a home remote control system. Some information is not clear at all when announced via screen reader; the UI arrangement is not particularly useful for providing the required data and simplifying interaction. However, certain functionalities offered by a HAS system could be very useful for a blind person in handling home activities (e.g., a button to set on/off lighting or open/close shutters). For overcoming observed limits and issues, some suggestions have been proposed to support developers of home automation systems to create more usable products.

Our analysis is limited to the Fibaro web interfaces; it would be extremely valuable to extend this investigation to other systems and to the interaction on mobile devices (gesture-based). Future work will further investigate the needs and expectations of blind people in order to better customize the interface to empower the user by improving their experience.

References

1. Al-Shaqi, R., Mourshed, M., Rezgui, Y.: Progress in ambient assisted systems for independent living by the elderly. SpringerPlus **5**(1), 624 (2016)
2. Brady, E., Morris, M.R., Zhong, Y., White, S., Bigham, J.P.: Visual challenges in the everyday lives of blind people. In: Proceedings of SIGCHI Conference on Human Factors in Computing Systems, pp. 2117–2126. ACM (2013)
3. Domingo, M.C.: An overview of the Internet of Things for people with disabilities. J. Netw. Comput. Appl. **35**(2), 584–596 (2012)
4. Javale, D., Mohsin, M., Nandanwar, S., Shingate, M.: Home automation and security system using Android ADK. Int. J. Electron. Commun. Comput. Technol.(IJECCT) **3**(2), 382–385 (2013)
5. Leporini, B., Paternò, F.: Applying web usability criteria for vision-impaired users: does it really improve task performance? Int. J. Hum.-Comput. Interact. **24**(1), 17–47 (2008)
6. Noury, N., Virone, G., Barralon, P., Ye, J., Rialle, V., Demongeot, J.: New trends in health smart homes. In: Proceedings of 5th International Workshop on Enterprise Networking and Computing in Healthcare Industry, Healthcom 2003, pp. 6–7, June 2003
7. Portet, F., Vacher, M., Golanski, C., Roux, C., Meillon, B.: Design and evaluation of a smart home voice interface for the elderly: acceptability and objection aspects. Pers. Ubiquit. Comput. **17**(1), 127–144 (2013)
8. Sandweg, N., Hassenzahl, M., Kuhn, K.: Designing a telephone-based interface for a home automation system. Int. J. Hum.-Comput. Interact. **12**(3-4), 401–414 (2000)

9. Santos, V., Bartolomeu, P., Fonseca, J., Mota, A.: B-live-a home automation system for disabled and elderly people. In: International Symposium on Industrial Embedded System, SIES 2007, pp. 333–336. IEEE, July 2007
10. Stefanov, D.H., et al.: The smart house for older persons and persons with physical disabilities: structure technology arrangements and perspectives. IEEE Trans. Neural Syst. Rehabil. Eng. **12**(2), 228–250 (2004)
11. W3C Accessible Rich Internet Applications (WAI-ARIA) 1.0, Recommendation, 20 March 2014

ALMA: An Indoor Localization and Navigation System for the Elderly

Sara Comai(✉), Emanuele De Bernardi, Andrea Masciadri, Matteo Matteucci, Fabio Salice, and Fabio Veronese

Politecnico di Milano, Piazza L da Vinci, 32, 20133 Milan, Italy
{sara.comai, emanuele.bernardi, andrea.masciadri, matteo.matteucci, fabio.salice, fabio.veronese}@polimi.it

Abstract. ALMA is an integrated system to support autonomous mobility and orientation for the elderly and, more in general, for the person with cognitive and/or mobility impairments in indoor facilities. It is composed of a set of modules for navigation, localization, planning, and autonomous wheelchair mobility, interfacing with the user. In this paper, we describe in detail the localization and navigation modules that can guide the target user to a selected destination and report our experience in a large care- and assistance-giving facility.

Keywords: Elderly · Indoor localization · Indoor navigation

1 Introduction

The chance of moving through the environment autonomously is crucial for the physical and psychological well-being of a person. However, there are cases (e.g., among elderly people) for which this requires capabilities that the person does not possess, or does not possess anymore, due to ageing or illness. The ALMA project [2] has developed a system that supports users in safely moving in indoor environments, from their current location to other locations of interest, considering their particular needs and/or limitations, and taking into account the state of the environment (e.g., presence of local obstacles). Such personal navigation assistant (PNA) aims at supporting or enhancing the autonomous motion of people with different degrees of mobility and/or cognitive impairment. In this paper, we introduce the overall system and, in particular, we focus on two of its main modules: the indoor localization module that identifies the current position of the user in the environment and the user interface, consisting of a mobile application supporting the navigation of the user to reach the desired destination.

The design of effective user interfaces to present navigation guidance is one of the challenges of indoor navigation services. In the ALMA project we have designed and tested a possible solution targeted at persons with cognitive and/or mobility impairments. Given the current location of the user, provided by the localization service, the indoor map and the path to reach the destination, instructions supporting orientation of the user are provided. A pilot application, presenting different scenarios for both primary (elderly, rehabilitation patients) and secondary (caregivers) end-users has guided the whole project development from user requirements, to final testing.

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 82–91, 2018.
https://doi.org/10.1007/978-3-319-76111-4_9

The paper describes the results of the ALMA project related to localization and navigation and is organized as follows: in Sect. 2 we describe the overall system and its architecture; in Sect. 3 we review works related to indoor localization and indoor navigation systems; Sects. 4 and 5 describe more in detail the localization module and the user interface; Sect. 6 presents experimental results of the pilot application; finally, Sect. 7 draws the conclusions and presents future works.

2 Overview and Architecture of the System

Figure 1 describes the overall architecture of the system. The ALMA project tackles the issue of not being able to move autonomously or efficiently by combining a set of advanced hardware and software technologies into an integrated and modular system composed by:

1. An indoor *Localization Module* based on a network of low-cost/low-power Radio Frequency (RF) emitters, to provide room level localization of people and objects; this module will be explained in detail in Sect. 4;
2. A system of networked smart cameras providing *video-based monitoring* for accurate indoor and outdoor localization, environment monitoring, and situation assessment;
3. An *Intelligent Module* for the online *planning and scheduling* of users' paths and activities, matching users' specific needs with the state of the environment (e.g., considering accessibility issues) and of available resources (e.g., a);
4. The *Personal Navigation Assistant*, a mobile application providing a user-friendly interface to all ALMA functionalities, tailored to the specific user requirements and physical limitations (e.g., with multimodal input/output modalities); this module will be explained in detail in Sect. 5;
5. The *Personal Mobility Kit* that, connected to a commercial electric wheelchair, allows users to perform automatic or assisted navigation and to interact with the surrounding environment.

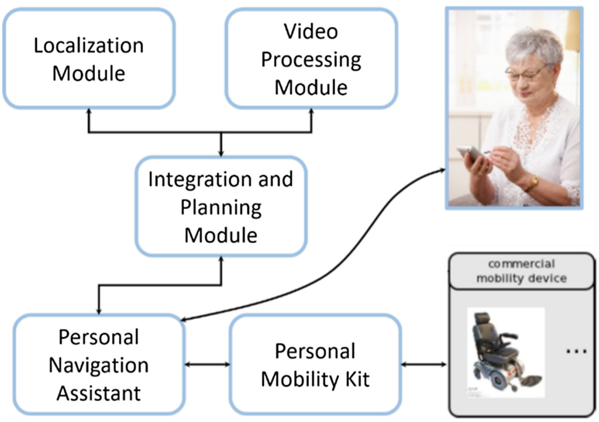


Fig. 1. Overview of the ALMA architecture

To support the different modules, a *Mapping Server* service collects and keeps updated a set of maps of the environment, providing information about

- physical obstacles, such as walls and fixed obstacles;
- terrain information, such as presence of steps/stairs or ease of accessibility for users with different needs and limitations, or for wheelchairs;
- passages connecting different areas, such as doors, and their current status (e.g., open/closed, unlocked/locked);
- possible destinations such as rooms, services or people, and their current status (e.g., if they are active/available);
- available services and their status (e.g., bathrooms, lifts).

ALMA users are therefore supported in their mobility to acquire knowledge about interesting locations (e.g., services, people, facilities, etc.), and to select and follow an efficient and safe path to such destinations considering their needs and/or limitations or the status of the environment. Secondary end-users, e.g., residences and hospitals, can leverage on the information collected by the system on the movements of their guests to monitor their aging, to design personalized support services and rehabilitation paths.

3 Related Works

Indoor Localization. Identifying, locating and tracking a target object within a specific environment is nowadays quite a common task. Several indoor localization technologies and approaches have been proposed, e.g., based on Radio Frequency (RF), acoustic waves, magnetic fields, etc. Lateration approaches allow to compute the current position; Dead Reckoning techniques continuously update the position (sometimes accomplished by speed and acceleration) of the target, given a previous position. Surely one of the most interesting and populated category is the RF-based indoor localization system, typically composed by a fixed network of beacons and a mobile worn device.

Dian et al. [6] present a WSN and RFID based localization system for indoor humans tracking, based on 802.11 active RF tags. They leverage the change in RSSI and the clustering of reference devices to obtain an improved localization accuracy up to 0.7 m–0.45 m, requiring however a double grid of devices.

Ubisense (Ubisense, UK 2005) [15] is a commercial location estimation system based on Ultra WideBand: RF signals are exchanged between active devices fixed in known locations, and those to be tracked. It can cover large areas and can manage a high number of users at the same time. Though quite accurate, the system can be challenging to install, due to the deployment of a timing cable, necessary for each device, and it has a high cost.

Radar (Microsoft 2000) [4] is probably the first example of positioning system using Wi-Fi networks providing WLAN (Wireless Local Area Network) access; a mobile device collects RSSI (Received Signal Strength Indicator) measurements from Wi-Fi access points and a fingerprinting algorithm is used to reconstruct the user position. In particular, instantaneous data are compared with a radio map of the environment to search for the most likely position. However, environment changes (humidity, furniture placement, etc.) can significantly affect the estimation process. The reported accuracy is 4.3 m, while the precision is 50% within 3 m. Many other systems based on WLANs have been later proposed (for example, AeroScout Company 2011, RFTechnologies 2011, from Deak et al. [5]), but they have similar characteristics.

User Interfaces for the elderly. Different communication channels and impairments of the elderly have been considered in designing user interfaces [8, 14].

Among the elderly, loss of vision is a major health care problem; approximately one of three elderly persons has some form of vision-reducing eye disease by the age of 65. According to [8] the interface should help older users to find items easily and should keep their attention focused on them, i.e., with no overlapping windows or icons, and a simple layout based on clarity and consistency. Black and white interfaces are preferable as well as sans-serif fonts. Authors of [9, 11, 12] have also investigated the best button size on a touchscreen interface for older adults: a size of 16.51 mm^2 is acceptable; buttons should be separated by a space of 3.17 mm to 12.7 mm. The universal font used for the texts in their applications is Helvetica, a basic sans-serif font, which could benefit older adults. The font sizes used in the design vary from 26 to 72 points.

According to [11] touchscreen mobile interfaces are preferred and not too difficult to use by the elderly. They suggested that larger buttons, icons and links (at least 8 mm) should be used, the gap between intended and actual touch location should be addressed, drag and pinch gestures should be considered rather than taps, and that multi-mode/view interfaces should be avoided.

Dexterity limitations (the effect of age-related degenerative changes in the musculoskeletal, vascular, and nervous systems) may affect the person's ability to manipulate objects and/or use arms, hands, or fingers. These issues require simple interfaces, big buttons, simple gestures [14]. Also hearing loss has to considered [13].

Indoor Navigation. Among the systems for indoor navigation, different proposals assist visually impaired users [17]; general systems are surveyed in [7]. Accurate and/or safe navigation of users with impairments towards a given point of interest in large and complex environments has been considered in different works. In particular, [10] focuses on the elderly and provides a system without wearable modules that exploits recognizable aids and abstractions to ensure that the elderly users can feel comfortable with it. The work presented in [3] proposes an Android-based navigation system for elderly people in hospital, based on RSSI and Dead Reckoning, reaching an accuracy of 93.3% but few details about the technique and the experiment are described. The authors in [1] developed a robotic wheeled walker able to assist people with moderate cognitive problems to navigate in complex indoor environments.

4 Localization Module

The ALMA Localization Module (LM) is a radio-based localization system used to track people (and objects), providing low computation, continuous recalibrated localization with uncertainty estimation. It is used by the ALMA Integration Module (IM) as one of the key sources of information required to assist its users: in particular, localization data are used to match user requirements with available resources (e.g., a nearby lift) and to manage potential difficulties (such as areas where the terrain is difficult for some categories of users). It is also used by the Personal Navigation System to assist the user in the navigation. The LM communicates with the IM in order to receive maps to be stored locally (maps contain also the positions of the RF beacons) and communicates the results of the localization (position and expected error) to the IM every 1000 ms.

One especially challenging problem is that of localizing moving targets: as most of the targets considered by project belong to this category, this problem has been given special attention. To solve it, a wireless sensor network with dynamic association has been used in the LM to retrieve and exploit RF signals for localization.

The localization algorithm leverages the possibility to update adaptively the relation between log-distance and RSSI. This can be achieved transmitting anchor-to-anchor signals and collecting the related RSSI values. Once every 20 s these values, collected by the server contribute to the update of an anchor-to-anchor signal matrix, using an autoregressive update model. Such matrix enables to build relations between the power values received by the mobile device and their distances through such a zero-calibration method [16].

Multilateration is implemented based on a discrete method DISCS (DIscrete Search for Candidate Solutions), exploring a grid of candidates solution, searching the most fitting point given the distances estimation. This enables on one hand the reduction of computation times still achieving an error comparable to gradient descent or particle filtering methods, and on the other hand it avoids wall-crossing and stabilizes the person position in the space. These has been proved with tests on the field as displayed in Figs. 2 and 3.

The performance of the LM in localizing targets on the map is compatible with the requirement of ALMA room-level localization. Precisely, targets are localized with a

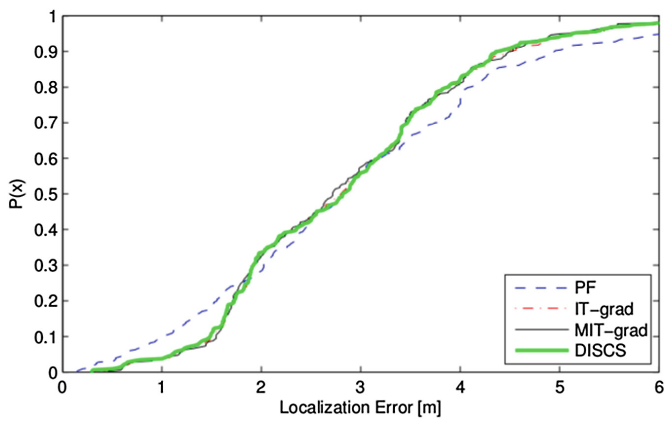


Fig. 2. Estimated Cumulative Density Function of Localization Error, comparison among DISCS, Iteration Terminated gradient descent (IT-grad), Magnitude of residual and Iteration Terminated gradient descent (MIT-grad), and Particle Filtering (PF)

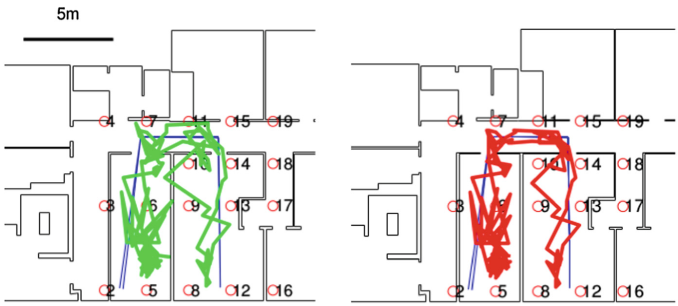


Fig. 3. Resulting localized path before (left, green) and after (right, red) the wall-crossing avoidance filtering. In blue the actual path. (Color figure online)

Circular Error Probable (CEP) of less than 4 m even with the lowest fixed devices density (0.04 m^{-2}) and this is compatible with the requirements of the specific application.

5 Personal Navigation Assistant

The Personal Navigation Assistant (PNA) is one of the key elements of the ALMA system: it a standard wireless device (a smartphone or a tablet) equipped with a software application, which enables the users to navigate through complex spaces and reach their destinations. The PNA has been designed for patients who need to autonomously move within a complex building (e.g., a hospital or a healthcare center): the PNA can guide them by constantly providing information about the direction to take, it can remind them the tasks of the day, and it can be a useful mean in case of help need. The user interface of the PNA is tailored to the target user and therefore it is easy-to-use, but together with the other modules of the ALMA system, it provides complex functionalities, such as: planning of routes that take into account the actual abilities and requirements of the user; monitoring of the progress of the user to detect possible deviations; possible re-planning of the routes to reach the desired destination. For instance, when local obstacles that require redirections are detected by the system.

The market already proposes several devices that can be turned into a PNA. However, the requirements of target users require simplified interfaces and the integration of specific functionalities that must be transparent to the user to enable patients/residents to reach their destinations independently and with safety. The proposed interface supports indoor navigation, manages a user's personalized agenda, and simple interactions in case of help request. The whole complexity is hidden to the final user, it is moved behind the interface and managed by the interaction between the PNA and the ALMA system.

To support navigation, the system exploits the map received from the ALMA server and gets a path computed by the integration module: the path is a sequence of points, which is processed to identify the doors to pass through, the turning points, and other elements such as lifts, stairs, etc. Figure 4 shows an example of a processed path. The

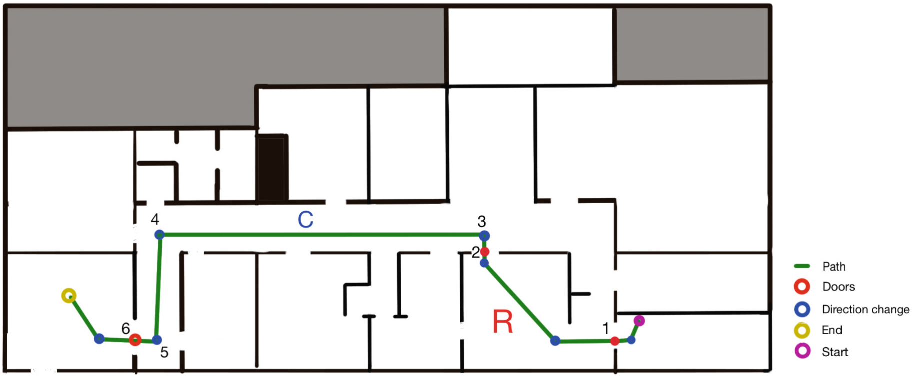


Fig. 4. Example of path segments

intermediate blue points indicate the direction change and the red points represent doors. They identify segments of the path. Depending on the type of the ending point of a segment, different instructions are provided.

Input/output multimodality is supported: users can use touch or voice input and receive information with text/images or audio. The modality may depend on the functionality that is used: e.g., to select a destination or browse the agenda, touch input is a feasible modality; instead, during navigation, audio output was requested to avoid the user focusing on the screen. Simple commands specifying the direction and exploiting landmarks such as doors, lifts, corridors etc. were used. For testing purposes a more advanced interface including also the map is available. Some snapshots of the PNA are shown in Fig. 5.

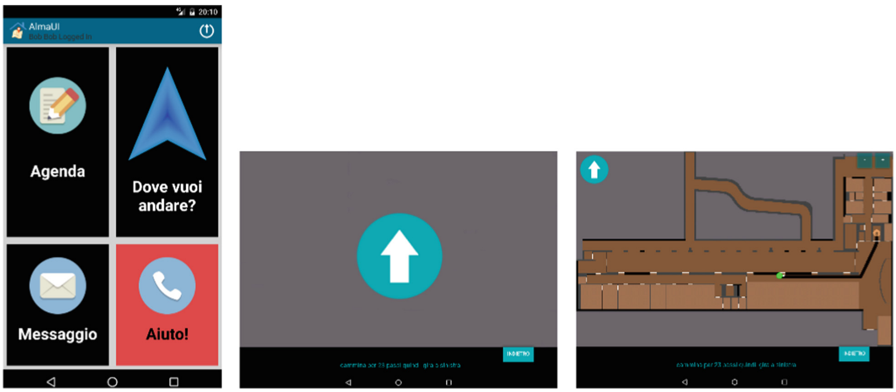


Fig. 5. (a) Main menu of the ALMA app with: agenda, navigation, messages, and help; (b) simplified navigation interface for target users; (c) advanced navigation interface used for testing.

The navigation system takes into account also possible *deviations*: they occur when the user during path progression is in an unwanted status or direction. Therefore, a proper reaction is anticipated in order to avoid her getting lost in the path. This is really important, especially for people with mobility or cognitive limitations. We have defined three different categories of deviations:

- **Positional Deviation:** when the user is in a place outside of the current path. For example, a user is inside the wrong room or takes a different corridor with respect to the planned one. In such a case the reaction consists in regenerating a new path by asking it to the server.
- **Temporal Deviation:** if there is a predefined maximum time to reach the destination (e.g., in case of a visit with a doctor at a given time) and the user may not be on time to reach the scheduled place, in this case s/he (as well as the staff) is notified about the delay.
- **Orientational Deviation:** when the user is located in the proper place, but he or she goes towards a wrong direction, which may lead her to go outside of the path or waste his/her time in the path progression. This may lead to the other two types of deviations, and the reaction toward consists in bringing the user back to the correct orientation.

6 Experimental Session

We performed some experiments in a care-giving facility. The Elderly Group consisted of n = 5 participants, who were in average 86.3 years old (SD 2.045), four females and one male. The included participants have similar physical and cognitive impairments, but they do not have strong visual or hearing impairments. The residents are at the facility on average since approximately 5 years, because of mild cognitive deficit (dementia) with memory disturbances. The level of autonomy in the activities of daily living is quite low and they need nursing assistance for simple activities, like washing. Most of the participant uses walking aids for displacements and walking (rollators, crunches). Regarding technology, they never use smartphones or tablets.

The path that the users had to follow included some turns and an unknown final destination (the kitchen elevator used only by cooking personnel).

Navigation commands were given using audio output. Because of the particularly noisy environment, the audio was enforced with a Bluetooth external speaker. We annotated comments made by the participants during the walk.

Most of the feedbacks were associated with the volume of the vocal commands, considered too low, despite the use of the external speaker. The interface was considered intuitive and the directions for the navigation were clear and easy to understand. The acceptance and the interest towards new technologies were rather high. During the displacement, the information given by the device did not affect the concentration and the stability of the participants.

All the users understood the commands and could press the correct buttons on the interface, with some problems with a left-handed user. The way they press the button depends on the person: in some cases, pressure was high and for an extended time; in other cases normal pressure was used; selection worked correctly in all the cases.

At the beginning of the navigation, they gave the impression to feel a little bit unsure, but immediately after, they continued on the path showing more confidence. The end of the path in front of a lift they have never used created some surprise to the participants since it was an unexpected destination: however, all of them could reach the selected destination.

7 Conclusions and Future Work

In this paper we have described the ALMA system targeted at elderly and, more in general, persons with cognitive and mobility impairments, to support indoor navigation. In particular, the system comprises a localization module and a personal navigation assistant application to guide the user to a selected destination.

Experimental results are promising for using the PNA as an aid in the rehabilitation process: the staff of the care facility believes that the PNA can be a useful auxiliary means to compensate for the cognitive deficits, but during the move the caregiver should be present anyway. More specific studies in this direction should be conducted.

Also the PNA may be further improved and include also user's adaptation of the navigation: by analyzing the path (e.g., in terms of speed, deviations, etc.) in relation with the characteristics of the user's profile, it should be possible to improve the whole system, as well as the PNA. For example, if by looking through the history or logging data of the user's paths we understand that following a specific part of a path is much harder for specific users or users with special restrictions, this should be considered.

Moreover, the LM can be improved by implementing more advanced power saving features, to reduce the maintenance costs and the user effort to keep the personal device recharged and operative. Finally, it would be interesting to implement dependability features among the different ALMA submodules, implementing methodologies based on redundant or correlated information [18].

Acknowledgments. This research is funded by AAL (Active and Assisted Living) Programme.

References

1. Aggravi, M., et al.: A smart walking assistant for safe navigation in complex indoor environments. In: Andò, B., Siciliano, P., Marletta, V., Monteriù, A. (eds.) Ambient Assisted Living. Biosystems & Biorobotics, vol. 11. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18374-9_45
2. ALMA (Ageing without Losing Mobility and Autonomy) Project. http://www.aal-europe.eu/projects/alma/
3. Aoki, R., Yamamoto, H., Yamazaki, K.: Android-based navigation system for elderly people in hospital. In: 2014 16th International Conference on Advanced Communication Technology (ICACT), pp. 371–377. IEEE, February 2014

4. Bahl, P., Padmanabhan. V.N.: RADAR: an in-building RF-based user location and tracking system. In: Proceedings of 19th Annual Joint Conference on IEEE Computer and Communications Societies, INFOCOM 2000, vol. 2. IEEE (2000)
5. Deak, G., Curran, K., Condell, J.: A survey of active and passive indoor localisation systems. Comput. Commun. **35**(16), 1939–1954 (2012)
6. Dian, Z., Kezhong, L., Rui, M.: A precise RFID indoor localization system with sensor network assistance. China Commun. **12**(4), 13–22 (2015)
7. Fallah, N., Apostolopoulos, I., Bekris, K.E., Folmer, E.: Indoor human navigation systems: a survey. Interact. Comput. **25**(1), 21–33 (2013)
8. Hawthorn, D.: Possible implications of ageing for interface designers. Interact. Comput. **12**, 507–528 (2000)
9. Jin, Z.X., Plocher, T., Kiff, L.: Touch screen user interfaces for older adults: button size and spacing. In: Stephanidis, C. (ed.) UAHCI 2007. LNCS, vol. 4554, pp. 933–941. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73279-2_104
10. Källström, M., Berdal, S., Joshi, S.G.: Designing an indoor navigation system for elderly people's capabilities. HCI **26**, 435–445 (2015)
11. Kobayashi, M., Hiyama, A., Miura, T., Asakawa, C., Hirose, M., Ifukube, T.: Elderly user evaluation of mobile touchscreen interactions. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011. LNCS, vol. 6946, pp. 83–99. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23774-4_9. https://dl.acm.org/citation.cfm?id=2042065
12. Murata, A., Iwase, H.: Usability of touchpanel interfaces for older adults. Hum. Factors **46**, 767–776 (2005)
13. Nunes, F., Silva, P.A., Abrantes, F.: Human-computer interaction and the older adult: an example using user research and personas. In: Proceedings of 3rd International Conference on Pervasive Technologies Related to Assistive Environments, PETRA, Samos, Greece (2010)
14. Phiriyapokanon T.: Is a big button interface enough for elderly uesrs? Toward user interface guidelines for elderly users. Master's thesis, Computer Engineering, Mälardalen University, Västerås (2011)
15. Steggles, P., Gschwind, S.: The Ubisense smart space platform. In: Adjunct Proceedings of 3rd International Conference on Pervasive Computing, vol. 191, pp. 73–76 (2005)
16. Redondi A., Tagliasacchi, M., Cesana, M., Borsani, L., Tarrio, P., Salice, F.: LAURA—LocAlization and Ubiquitous monitoRing of pAtients for health care support. In: 2010 IEEE 21st International Symposium on Personal, Indoor and Mobile Radio Communications Workshops (PIMRC Workshops), pp. 218–222. IEEE (2010). http://ieeexplore.ieee.org/document/5670365/
17. Tsirmpas, C., Rompas, A., Fokou, O., Koutsouris, D.: An indoor navigation system for visually impaired and elderly people based on Radio Frequency Identification (RFID). Inf. Sci. **320**, 288–305 (2015)
18. Veronese, F., Comai, S., Salice, F.: Method, design and implementation of an indoor tracking system with concurrent fault localization. EAI Endors. Trans. Ubiquitous Environ. **15**(4), e3 (2014)

Exergames in Individuals with Down Syndrome: A Performance Comparison Between Children and Adolescents

Adrian Macias[1], Karina Caro[2], Luis A. Castro[1](✉), Veronica Sierra[1], Edgar A. Ahumada[1], and Iván A. Encinas[1]

[1] Sonora Institute of Technology (ITSON), Ciudad Obregon, Sonora, Mexico
{adrian.macias, veronica.sierra}@itson.edu.mx, luis.castro@acm.org, eahumadasolorza@gmail.com, ivanencinasm@gmail.com
[2] Drexel University, Philadelphia, PA, USA
karinacaro@drexel.edu

Abstract. Individuals with Down Syndrome (DS) have delays in the development of motor function associated with impairments including difficulty with precise movements of limbs, poor balance, and poor visual-motor coordination. It has been reported that children and adolescents with DS might present differences in terms of visual-motor coordination skills, task persistence, emotional expressions, among others. Exergames have the potential to support motor coordination as they combine physical exercise with gaming technology. However, little has been said about the game experience of individuals with DS playing exergames. This work presents the results of an exploratory study of 10 individuals with DS playing a commercial exergame. Our results show a significant difference between children and adolescents in terms of task-efficacy, selective attention, and prompts. Finally, we discuss our results and the implications for designing exergames to support motor coordination of people with DS.

Keywords: Exergames · Kinect Monsters · Motor coordination
Down Syndrome

1 Introduction

Down Syndrome (DS) is a chromosomal disorder caused by the presence of an extra chromosome, occurring in approximately 1 in 700 births [1]. Children with DS show relative delays when compared to typically developing children of the same chronological age, move through stages of development more slowly, and exhibit more within-group variability than typically developing children [2]. Particularly, children with DS have delays in development of motor function associated with impairments including difficulty with precise movements of limbs (e.g., stepping over a stick while on a balance beam), low muscle tone, poor postural control, poor balance, poor visual-motor coordination and poor visual-spatial perception skills, and, for some children, congenital heart disease, and obesity [3]. These deficits in motor development might be

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 92–101, 2018.
https://doi.org/10.1007/978-3-319-76111-4_10

of particular importance for occupational performance in school, daily living, play, and performance in other areas of occupation [4].

It has been demonstrated that play is an important facilitator and an important means for the development of cognitive and motor skills [5]. In addition, several studies have shown that videogames are adequate tools for supporting skills in different areas of development (e.g., cognitive, social, motor) [6]. Exergames, videogames that require physical exercise by players as input to control the game, have the potential to support motor skills of different populations such as older adults [7] and individuals with cerebral palsy [8]. However, little has been said about the player experience of individuals with DS playing exergames to support motor coordination. In this work, we aim at exploring the use of a commercially available exergame in individuals with DS. In particular, we aim at understanding the differences between children and adolescents with DS, since it has been reported that young children with DS show lower levels of task persistence and higher levels of off-task behavior, interfering with task completion, when compared to adolescents or young adults with DS [9]. On the other hand, in [9, 10], they show that young children with DS are more likely to ask for help, when faced challenging tasks, compared with typically developing children.

Additionally, other studies have found relative strengths in visual-spatial perception skills such as visual memory and visual-motor coordination in older children and young adults when compared to other skills (e.g., verbal processing skills) [11]. In terms of emotional expressions, studies have shown that older individuals with DS show relative strengths in nonverbal communication (e.g., facial expressions, gestures and eye contact) [12]. However, other studies have shown that children with DS may also send more positive emotional signals than other children with mental retardation. In one study, 5- to 12-year-old children with DS smiled more frequently than children with other intellectual disabilities, although smile frequency changed as individuals with DS approached adulthood [13]. Although there are studies that show that individuals with DS commonly play videogames (e.g., 90% of 21 individuals with DS [14]), differences in game experience between children and adolescents with DS have not been extensively reported. A study with five children with DS between 6 and 12 years old [15] shows that older children ($\sim$12 years old) found some scenarios very childish when compared to young children.

In this work, ten individuals with DS, 5 children and 5 adolescents, played a Kinect-based exergame designed to support motor coordination skills such as visual-motor coordination and visual-spatial perception. Our results show a significant difference between children and adolescents in terms of task efficacy (i.e., the rate of successful attempts to complete the tasks asked by the exergame), selective attention (i.e., the capability to focus on the exergame ignoring competing distractions), and prompts required to play the exergame. No significant differences were found in children and adults in terms of on-task behavior (i.e., the capability to focus on the exergame), positive emotional expressions, and game experience. These results show that exergames might have the potential to support motor coordination of individuals with DS, but important considerations should be made in terms of the design of exergames for children or adolescents with DS.

2 Related Work

There are several studies that use exergames for supporting motor coordination of individuals with special needs such as older adults [16], cerebral palsy [17] and autism [18]. Few studies have investigated the use of exergames in individuals with DS. A systematic review of the use of exergames in children with neuromotor dysfunction [19] presents only two studies including individuals with DS, both used the Wii gaming technology. The work of [20] presents the effectiveness of a virtual reality (VR) system using the Wii gaming technology when compared to traditional therapy. 105 children between 7 and 12 years old were assigned to Wii-based therapy and 50 to traditional therapy. The results suggest that VR using the Wii gaming technology shows benefit in improving sensorimotor functions among children with DS when compared to traditional therapy. In [21], it is shown that the efficacy of a 2-month Wii-based therapy in 18- to 60-year-old adults with DS. The results suggest that Wii-based therapy can be an effective tool to improve physical fitness, functional mobility, and motor proficiency in adults with DS, including crucial measures such as aerobic capacity and lower limb strength. Also, in [22], the authors present a virtual system for upper limbs rehabilitation in children using a haptic device and Oculus Rift. Using the system, children perform hand rehabilitation exercises such as extension and flexion of wrist and finger pinch. The system was evaluated through a usability study with four children with DS (between 9 and 12 years old). The system was positively rated for rehabilitation, felt immersed, and enjoyed the game.

The related work shows the outcomes of using exergames based on Wii gaming technology and haptic devices in children and adults with DS. However, to the best of our knowledge, little has been said about the use of exergames that use only body movements (e.g., upper-limb movements) beyond fingers on handheld devices (e.g., Wii mote) such as Kinect-based exergames in individuals with DS. Kinect-based exergames could elicit a different player experience in this population. Although several studies have shown differences between children and adolescents with DS in terms of task persistence, off-task behavior, level of required help to complete a task, among others [9–12], the literature shows that none of the studies presents differences between children and adolescents with DS playing exergames.

3 Methods

We have been working closely with an institute for intervention and education for individuals with DS, partially funded by a non-profit NGO, located in a mid-size city of Northwest, Mexico. We carried out a quasi-experiment, in which ten individuals with DS were recruited to play an exergame in a controlled setting.

Participants. We recruited 10 individuals with DS between 6 and 14 years old (AVG age = 10.9 years, ±2.9 years, 5 males), including 5 children (AVG age = 8.4 years, ±1.7 years, 2 males) and 5 adolescents (AVG age = 13.4 years, ±0.9 years, 3 males). Nine of the participants had previous experience with computer games such as puzzles and crosswords. However, only one participant had previous experience playing exergames (i.e., Kinect adventures).

Variables. The variables we are interested in studying are:

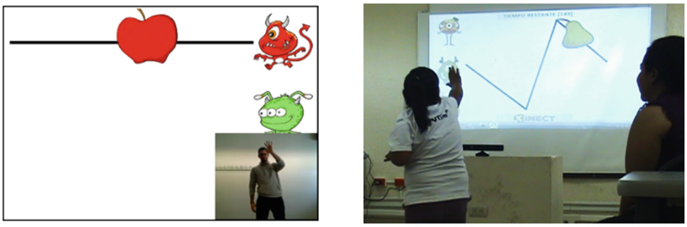
– *Independent variables.* Chronological age: Children (<=10) vs. Adolescents (>10)
– *Dependent variables.* Efficacy (%: successful attempts/total attempts), emotional expressions (%: positive expressions time/total time), attention (%: on-task time/total time), on-task attention (%: selective attention time/on-task attention total time), prompts (%: prompting time/total time), Game experience (N: ratings of perceived game experience through the Game Experience Questionnaire [23]).

Hypotheses. Based on previous knowledge from the literature, we propose six hypotheses we are interested in. H1 to H5 are related to differences in children and adolescents with DS in terms of task persistence [9] (H1), off-task behavior [9, 11] (H3, H4), level of required help to complete a task [9, 10] (H5) and emotional expression [12, 13] (H2). H6 is related to the experience of individuals with DS playing videogames [14, 15], specifically exergames [20–22], as it has been reported that individuals with DS have successfully played videogames and Wii-based exergames, but differences between children and adolescents with DS when playing videogames and exergames have yet to be studied.

– **H1.** *Efficacy:* The proportion of successful attempts will be significantly higher in adolescents than in children with DS.
– **H2.** *Emotional expressions:* The proportion of positive emotional expressions will be significantly higher in children than in adolescents with DS.
– **H3.** *Attention:* The proportion of time for on-task attention will be significantly higher in adolescents than in children with DS.
– **H4.** *On-task attention:* The proportion of time for selective attention will be significantly higher in adolescents than in children with DS.
– **H5.** *Prompts:* The proportion of time spent on prompting will be significantly higher in children than in adolescents with DS.
– **H6.** *Game experience:* There will be no difference in the perceived rating of game experience between children and adolescents with DS.

3.1 Instruments

Kinect Monsters exergame. We used the commercially-available exergame Kinect Monsters (KM). KM is a videogame to support neuro-rehabilitation, mainly for children. Players must feed monsters following a black line (i.e., pattern) (see Fig. 1). Using their hand, participants can move and pick up a fruit by positioning the cursor over it. Once the fruit has been picked up, the cursor fades away, and the fruit is directly controlled by the player's hand. If the player is unable to follow the line, the player must start again. Once the player has successfully taken the fruit to the monster, the word “Good” is shown, and another fruit is randomly shown. KM uses Kinect to track players' movements, thus no additional on-body sensors are needed.



(a) (b)

Fig. 1. (a) KM's interface. Source: kinectfordevelopers.com. (b) A participant playing KM.

Game Experience Questionnaire. To measure the game experience of our participants, we used the in-game module of the Game Experience Questionnaire (GEQ) [23]. The GEQ was answered by a proxy (therapist).

Screen recorders. All on-screen events were recorded for task completion analysis (i.e., successful vs. failed attempts).

Video cameras. Two video cameras (front and rear) were used to record participants.

BORIS. We used BORIS [24] to facilitate video coding and analysis.

3.2 Research Procedure

During the sessions, a therapist and the participants' parents were present at all time. The configuration of data collection was as follows:

- **Stage 1: Pre-play (5 min).** We collected demographics and consent forms signed by the participants' parents.
- **Stage 2: Play (~5 min).** The therapist instructed participants on how to play KM, before proceeding to play. Instructions to participants included how they should move their arms, where to step on (i.e., we drew a square on the floor), open their hands to grab a fruit, and the like. The aim was to feed 5 monsters or play for 5 min. non-stop, whatever happened first. If a participant was unable to complete at least one task (e.g., feed a monster) within 5 min., the game was stopped.
- **Stage 3: GEQ (5 min).** The GEQ was applied to each participant in a structured fashion, i.e., the therapist was asking the questions and filling out the electronic forms designed for collecting data. In the case of low functioning students, the therapist worked as a proxy answering the GEQ; otherwise, participants answered it themselves.

3.3 Data Collection

Participants' performance was collected by non-participant observation during the exergame sessions. In each session, three members of the research team were involved, one therapist, and the participants. The therapist was the only one who interacted with the participants. The sessions were carried out at the lab, as per request of the institute we are collaborating with. Total effective game play time collected includes: Front-facing camera (42:31 min.) and rear-facing camera (42:30 min.).

3.4 Data Analysis

For the analysis, we created a coding scheme based on the literature that included prompts: physical, verbal, gestural, and modeling, or a combination of any of those; emotional expressions: positive and negative [25]; attention: on-task (selective, divided), and off-task [26]; attempt outcome: successful or failed. The coding process was independently carried out by two of the authors. Before coding the whole set of videos, coders refined the schema and criteria with a pilot test in the same conditions. This pilot coding session was collaboratively carried out by the coders. In the end, all differences in coding were discussed by the coders. Videos were replayed as needed. Coders compared their initial set of codes to the final (agreed) one. Accuracy was: 99.6, 94.0, 84.5% for attention, prompts, and emotional expressions respectively for the first coder; and 95.9, 86.7, 78.2% for the second coder. Aggregate accuracy per coder was 94.9% for the first, and 89.6% for the second one.

To determine if there is a significant difference between children and adolescents, we used an independent *t* test, since our data presented a parametric distribution. The normality of the data was calculated using a Shapiro-Wilk test. Equality of variances was not assumed in the data.

4 Results

In this section, we present the results obtained in the experiments. Our results are organized by dependent variables. Table 1 shows a summary of our results.

Efficacy. Efficacy was the proportion of successful attempts per participant. Some participants performed all five attempts, while some other managed to perform fewer attempts. The proportion of successful attempts was significantly higher for adolescents (mean = 0.93) than for children (mean = 0.52). The mean difference was 0.41. An independent *t*-test showed that the difference between conditions was significant ($t = -2.7608$, $df = 5.8317$, *p*-value = 0.0169, one-tailed). Thus, H1 was accepted.

Table 1. Performance by participants by dependent variables

	Children			Adolescents			Sig.
	AVG (SD)	Min	Max	AVG (SD)	Min	Max	
Successful attempts (%)	0.52 (0.30)	0.00	0.75	0.93 (0.15)	0.67	1.00	*
Pos. emo. expressions (%)	0.20 (0.19)	0.01	0.48	0.03 (0.03)	0.00	0.08	
Attention (%)	0.93 (0.11)	0.73	100	1.00 (0.00)	1.00	1.00	
Selective attention (%)	0.82 (0.04)	0.77	0.86	0.97 (0.4)	0.90	1.00	***
Prompts (%)	0.23 (0.15)	0.11	0.48	0.08 (0.05)	0.03	0.16	*
Game exp. (score)	2.09 (0.49)	1.36	2.64	2.27 (0.33)	1.86	2.64	

Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Emotional expression. We computed the aggregated time spent in which positive emotional expressions were shown by the participants and divided by the total time of the session per participant. The percentage of time in positive emotional expressions was higher for children (20%, SD = 19%) than for adolescents (3%, SD = 3%). The mean difference was 17%. An independent *t*-test showed that the difference between conditions was not significant ($t = 1.9679$, $df = 4.2469$, p-value = 0.05817, one-tailed). Thus, H2 was rejected.

Attention. We computed the total time in focusing participants' attention on the requested task (i.e., playing the exergame) vs. being off-task. Adolescents were able to maintain 100% (SD = 0%) their attention on-task, while the percentage of children was lower (93%, SD = 11%). The mean difference was 7%. An independent *t*-test showed that the difference between conditions was not significant ($t = -1.4486$, $df = 4$, p-value = 0.1105, one-tailed). Thus, H3 was not supported by our data.

Selective attention. Attention on-task was divided into selective attention and divided attention. Selective attention means that participants can be completely focused on the requested task. The proportion of time spent in selective attention was significantly higher for adolescents (97%, SD = 4%) than for children (82%, SD = 4%). The mean difference was 15%. An independent *t*-test showed that the difference between conditions was significant ($t = -5.4385$, $df = 7.9819$, p-value = 0.0003109, one-tailed). Thus, H4 was accepted.

Prompts. The proportion of time in which the therapist was prompting the participants was computed per participant. The proportion of prompts was significantly higher for children (23%, SD = 15%) than for adolescents (8%, SD = 5%). The mean difference was 15%. An independent *t*-test showed that the difference between conditions was significant ($t = -2.0539$, $df = 5.0207$, p-value = 0.04747, one-tailed). Thus, H5 was accepted.

Game experience. In Table 2, we present the seven dimensions of the GEQ. The perceived game experience score obtained was higher for adolescents (2.27, SD = 0.33) than for children (2.09, SD = 0.49). The mean difference was 0.18. An independent

Table 2. Average and standard deviation of the GEQ results by dimensions (Min = 0, Max = 4)

	Competence	Immersion	Flow	Tension	Challenge	(−) Affect	(+) Affect
Children							
AVG	2.90	3.00	0.60	0.30	**3.50**	**0.60**	3.70
SD	1.39	0.94	0.42	0.45	**0.87**	**0.89**	0.67
Adolescents							
AVG	**3.13**	**3.63**	**1.38**	**0.25**	3.25	0.25	**4.00**
SD	**0.85**	**0.48**	**0.25**	**0.50**	0.96	0.50	**0.00**
Overall							
AVG	3.29	3.17	0.83	0.28	3.22	0.56	3.83
SD	1.20	0.75	0.61	0.44	0.87	0.88	0.50

t-test showed that the difference between conditions was not significant ($t = 0.52345$, df = 5.6273, p-value = 0.6206, two-tailed). Thus, the null hypothesis H6 was accepted.

5 Discussion

Our results show that, for the purposes of education and therapies of any kind, there are aspects that need to be taken into consideration when using exergames with children and adolescents with DS, as significant differences were found in efficacy, selective attention, and prompting between them. Our results indicate that our participants who were older than 10 years were more focused on the task they were working on, which could also result in a higher proportion of successful attempts. On the other hand, our participants who were 10 years old or younger needed considerably more prompting by the therapist. Both of these aspects may have practical implications not only for the design of much more considerate exergames (e.g., if the player is a child, the exergame should provide different kinds of prompts such as visual and verbal ones, including a step-by-step guide when children face challenges during the gameplay), but also in the time invested in each of these children per session.

We found no difference in game experience between children and adolescents, meaning that, regardless the age, individuals with DS can have similar experiences. Regarding game experience, our adolescent participants felt more competent and skillful, which was understandable since their proportion of successful attempts was higher. Since adolescents were more focused on the tasks, they also reported higher levels of immersion than children, meaning that they were really interested in the game. Although slightly higher in children, tension or frustration was minimal for both groups, younger participants showed more positive expressions. Finally, positive and negative affect were generally perceived better for adolescents than for children, but no significant differences were found.

From our observations, those participants who were not able to complete the whole set of attempts did not show any sign of frustration or irritation when an audio was played warning of the end of the game. Also, when adolescents could not successfully complete an attempt, they looked for alternate ways to complete it. This was not the case of children, who often kept performing the same movements, even if the previous ones were not successful. The latter, we believe, could be related to the developmental age of the participants, which tend to be more developed in individuals with higher chronological age. Finally, in terms of the area (i.e., a drawn square) where participants were supposed to stand on, adolescents tended to stay within boundaries, whereas children were generally less aware of this area, resulting in challenging tracking. This could indicate that for the design of exergames for children with DS, tracking should consider that children might not be able to stay in the same area for long time intervals, tracking children's limbs, without needing to stay in one place.

The main limitation of our study was the sample size since it was rather small. Still, sample size is comparable to those found in the literature. We need much more controlled studies with a higher number of participants to present conclusive findings.

6 Conclusions and Future Work

We presented a study with individuals with DS, in which participants played an exergame in a controlled setting, accompanied by a therapist. Our results indicate that there are several aspects that merit careful consideration when working with this population. First, adolescents tend to be much more focused on the tasks, which yield higher proportion of successful attempts. Also, selective attention was significantly higher in adolescents than in children. Regarding therapist prompting, children seem to need more of those, meaning that this requires a presumably higher effort from the therapist. However, neither of those factors affected the overall game experience of the participants. Future work includes working closely with the institute for carrying much more focused experiments (e.g., different types of games), which can help them take decisions in terms of therapies and educational strategies.

Acknowledgements. We thank the participants, family members, and staff of CEART DOWN. This work was partially funded by a research grant by the Instituto Tecnólogico de Sonora and the PFCE federal program in Mexico.

References

1. Davis, A.S.: Children with Down syndrome: implications for assessment and intervention in the school. Sch. Psychol. Q. **23**, 271–281 (2008)
2. Spanò, M., Mercuri, E., Randò, T., Pantò, T., Gagliano, A., Henderson, S., Guzzetta, F.: Motor and perceptual-motor competence in children with Down syndrome: variation in performance with age. Eur. J. Paediatr. Neurol. **3**, 7–14 (1999)
3. Wang, H.Y., Long, I.M., Liu, M.F.: Relationships between task-oriented postural control and motor ability in children and adolescents with Down syndrome. Res. Dev. Disabil. **33**, 1792–1798 (2012)
4. Fidler, D.J., Hepburn, S.L., Mankin, G., Rogers, S.J.: Praxis skills in young children with Down syndrome, other developmental disabilities, and typically developing children. Am. J. Occup. Ther. **59**, 129–138 (2005)
5. Gallahue, D.L., Ozmun, J.C.: Understanding Motor Development Infants, Children, Adolescents, Adults. McGraw-Hill Humanities, Social Sciences & World Languages, New York (1998)
6. Thompson, D., Baranowski, T., Buday, R., Baranowski, J., Thompson, V., Jago, R., Griffith, M.J.: Serious video games for health: how behavioral science guided the development of a serious video game. Simul. Gaming. **41**, 587–606 (2010)
7. Gerling, K.M., Mandryk, R.L., Linehan, C.: Long-term use of motion-based video games in care home settings. In: CHI 2015, pp. 1573–1582 (2015)
8. Hernandez, H.A., Ketcheson, M., Schneider, A., Ye, Z., Fehlings, D., Switzer, L., Wright, V., Bursick, S.K., Richards, C., Graham, T.C.N.: Design and evaluation of a networked game to support social connection of youth with cerebral palsy. In: ASSETS 2014, pp. 161–168 (2014)
9. Pitcairn, T.K., Wishart, J.G.: Reaction of young children with Down's syndrome to an impossible task. Br. J. Dev. Psychol. **12**, 485–498 (1994)
10. Kasari, C., Freeman, S.F.: Task-related social behavior in children with Down syndrome. Am. J. Ment. Retard. **106**, 253–264 (2001)

11. Fidler, D.J.: The emerging Down syndrome behavioral phenotype in early childhood. Infants Young Child. **18**, 86–103 (2005)
12. Chapman, R.S.: Language development in children and adolescents with Down syndrome. Ment. Retard. Dev. Disabil. **3**, 307–312 (1997)
13. Fidler, D.J., Philofsky, A., Hepburn, S.L., Rogers, S.J.: Nonverbal requesting and problem-solving by toddlers with down syndrome. Am. J. Ment. Retard. **110**, 312–322 (2005)
14. Prena, K.: Down syndrome videogame preferences (2014)
15. Macedo, I., Trevisan, D.G., Vasconcelos, C.N., Clua, E.: Observed interaction in games for Down syndrome children. In: Hawaii International Conference on System Sciences, pp. 662–671 (2015)
16. Gerling, K.M., Miller, M., Mandryk, R.L., Birk, M.V., Smeddinck, J.D.: Effects of balancing for physical abilities on player performance, experience and self-esteem in exergames. In: CHI 2014, pp. 2201–2210 (2014)
17. Altanis, G., Boloudakis, M., Retalis, S., Nikou, N.: Children with motor impairments play a kinect learning game: first findings from a pilot case in an authentic classroom environment. Interact. Des. Archit. J. - IxD&A **19**, 91–104 (2013)
18. Caro, K., Tentori, M., Martinez-Garcia, A.I., Alvelais, M.: Using the FroggyBobby exergame to support eye-body coordination development of children with severe autism. Int. J. Hum. Comput. Stud. **105**, 12–27 (2017)
19. Hickman, R., Popescu, L., Manzanares, R., Morris, B., Lee, S.-P., Dufek, J.S.: Use of active video gaming in children with neuromotor dysfunction: a systematic review. Dev. Med. Child Neurol. **59**(9), 903–911 (2017)
20. Wuang, Y.P., Chiang, C.S., Su, C.Y., Wang, C.C.: Effectiveness of virtual reality using Wii gaming technology in children with Down syndrome. Res. Dev. Disabil. **32**, 312–321 (2011)
21. Silva, V., Campos, C., Sá, A., Cavadas, M., Pinto, J., Simões, P., Machado, S., Murillo-Rodriguez, E., Barbosa-Rocha, N.: Wii-based exercise program to improve physical fitness, motor proficiency and functional mobility in adults with Down syndrome. J. Intellect. Disabil. Res. **61**, 1–11 (2017)
22. Pruna, E., Acurio, A., Tigse, J., Escobar, I., Pilatásig, M., Pilatásig, P.: Virtual system for upper limbs rehabilitation in children. In: De Paolis, L.T., Bourdot, P., Mongelli, A. (eds.) AVR 2017. LNCS, vol. 10325, pp. 107–118. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60928-7_9
23. IJsselsteijn, W., Poels, K., de Kort, Y.A.: The Game Experience Questionnaire: Development of a Self-report Measure to Assess Player Experiences of Digital Games. TU Eindhoven, Eindhoven (2008)
24. Friard, O., Gamba, M.: BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods Ecol. Evol. **7**, 1324–1330 (2016)
25. Ekman, P., Friesen, W.V., Hager, J.C.: Facial Action Coding System - Investigator's Guide (2002)
26. Corbetta, M., Miezin, F., Dobmeyer, S., Gordon, L., Shulman, L., Petersen, S.: Selective and divided attention during visual discriminations shape, color, and speed: functional anatomy by positron emission tomography. J. Neurosci. **11**, 2393–2402 (1991)

Digital Invasions Within Cultural Heritage: Social Media and Crowdsourcing

Lorenzo Monti[1(✉)], Giovanni Delnevo[1], Silvia Mirri[2], Paola Salomoni[2], and Franco Callegati[1]

[1] Interdepartmental Centre for Industrial ICT Research - CIRI ICT, Università di Bologna, Bologna, Italy
lorenzo.monti20@unibo.it
[2] Department of Computer Science and Engineering, Università di Bologna, Bologna, Italy

Abstract. The wide diffusion of mobile devices and of digital technologies are dramatically changing the usage scenarios in different contexts. One of them is cultural heritage, where new media are offering huge possibilities for the enhancement and the enrichment of heritage experience, improving the users' involvement. In particular, tourists equipped with their mobile devices are invading cultural attractions, sharing pictures and comments (together with hashtags and geo-localized positions) on social networks. These represent a source of data, which can be integrated with the official ones provided by GLAM (Galleries, Libraries, Archives, and Museums) and cultural heritage institutions, enriching them. In this paper, we explore how social networks and crowdsourcing activities can be exploited as a source of information for cultural places and pieces of art.

Keywords: Web scraping · Data extraction · Open data
Cultural heritage · Crowdsourcing · Social network
Crowdsourced data · Social media data

1 Introduction

Society is moving toward the post-industrial age, where the consumer, as we know, no longer exists and s/he is replaced with a new kind of user that is not just a mere consumer, but also a producer of contents [1]. These new users are called *prosumer*, term branded by Toffler [2]. From this perspective, there will be new professionals and there will be an unavoidable metamorphosis from the actual ones, driven by the rise of resilient factors mainly from the self-adjustment to new job needs. The first half of twentieth century was certainly marked by a huge technological change. An example of this change is the advent of cinematography, that have certainly been an impact also in the way to interpret and experience cultural heritage. One of the most important dissertation about

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 102–111, 2018.
https://doi.org/10.1007/978-3-319-76111-4_11

this theme was written by Benjamin, in his famous work 'The Work of Art in the Age of Mechanical Reproduction' in [3]. He claims that the introduction of a new technique to produce, reproduce and spread worldwide artworks, has radically changed the attitude towards art of both artists and public [4]. The author focuses on the particular kind of cultural heritage, that is artwork which, through mechanical reproduction techniques like films rather than printing, destroyed the concept of "aura" of an artwork. The aura is broadly understood as the feeling of religious nature resulted in the spectator in front of the original specimen of an artwork. As with the advent of mechanical reproduction, new digital media have radically altered the concept of cultural heritage: digitization technique, animation, 3D reconstruction or immersive learning are only some examples and have certainly changed and redefined the concept of transmission and use of knowledge [5]. The use of multi-touch monitors used like information tools, the realization of consoles with reconstructions of three dimensional digital model, visualizations of active/passive anaglyph, immersive augmented reality, synchrony or diachrony information analysis or applications for low-vision users based on touch tactile system are sample scenarios of media that promote a different interaction with knowledge so as to improve visitors' engagement [6]. In Web applications, the progressive technological development has led to the creation of different platforms with expanded visualization and navigation capabilities of three-dimensional data more and more pertinent to the perceptual image quality and able to involve the user both in space and time terms [7].

New media offers enormous possibilities for the enhancement and enrichment of heritage experience and interpretation. In the urban context, municipalities and public entities should provide adequate infrastructure, so as to guarantee distribution and sharing of multimedia resources, such as audio and video [8]. The question is how to make best use of new media to maintain the integrity of heritage artifacts and sites. How this has to be achieved can vary according to particular heritage contexts, artifacts and sites, and it can also differ according to various curatorial practices and different media [9].

In these area lies the concept of *crowdsourcing*, a neologism used for the first time by Howe in [10]. It is formed by two words, *crowd* and *outsourcing* that indicates the outsourcing of data supply. There is no unique definition of crowdsourcing, but thanks to the work of Estellés-Arolas and González-Ladrón-De-Guevara [11], that tries to collect all the definitions of it and tries to identify the common characteristics of them, we can define as "a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit." Indeed, the increasing diffusion of mobile devices have contributed to the use of crowdsourcing also in geo-spatial context [12], changing the way the informations are produced, used and stored [13].

Tourists equipped with their smart phones are invading cultural attractions, taking pictures and selfies, sharing posts, hashtags [14], and their geo-localized position, tagging friends and places. In this way, they are populating and enriching social networks with such digital media (sometimes within specific events planned or supported by cultural institutions or by social media communities, such as the Wiki Loves Monuments[1] and the Digital Invasions[2] initiatives). These new contents could be exploited by cultural institutions, tourism offices, public entities, and private foundations, with the aim of monitoring activities related to their goods and places on social media. In particular, the use and the collection of pictures coming from social networks, from non-common devices (i.e. vehicles on board cameras [15]), from IoT and cloud contexts can play a significant role.

With this goal in mind, in this paper we present an experiment we have conducted, with the aim of gathering, grouping, and evaluating media and content coming from crowdsourcing activities in social media, related to cultural heritage (i.e. historical palazzos, monuments, museums, pieces of art, etc.), by using scraping techniques.

The remainder of the paper is structured as follows. Section 2 provides a brief description of the background and presents some related work. Section 3 summarizes the approach we have exploited. Section 4 describes the architecture and the implementation of the system, while the results are illustrated and discussed in Sect. 5. Finally, Sect. 6 concludes the paper highlighting some final remarks and future work.

2 Background and Related Work

Since the late 90s' the potential of new media has improved many areas and among them there cultural heritage. This has occurred in various forms. An example is the reconstruction, based on 3D modeling, of archaeological sites or lost monuments and cultural heritage due to war conflicts (just like Palmyra site in Syria) or due to natural disasters (just like tsunamis or earthquake [16]). Thanks to crowdsourcing and to motion technique, used for the reconstruction of a 3D model, it has been possible to reconstruct a model of Plaka Bridge, a 19th-century stone one-arch bridge in Greece, that collapsed during the floods of 1 February 2015. In this project, acquired images are uploaded from different kind of users, with the result that they have different resolution, perspective, distance and brightness. Thus, it has to be considered that they are also snapped in different seasons and times.

Researchers also investigated how transcription is even more important, in historical documents, than the uniqueness of most of these documents and the preservation of their contents is essential for historical and cultural reasons. Transcribing handwritten documents through OCR technologies isn't always enough and is far from perfect. Crowdsourcing and human experts' revision emerged as a powerful tool in order to obtain a correct transcription. Through the rise of a

[1] https://www.wikilovesmonuments.org/.
[2] http://www.invasionidigitali.it/en/.

crowdsourcing platform [18], in which users contribute to a small given task, getting back a little or null payment, the transcription of cultural texts has spread. We can find examples in tranScriptorium [19], an European project that aims to develop a solution for annotating handwritten historical documents using modern, holistic Handwritten Text Recognition (HTR) technology, or alternative solutions, such as the use of speech dictation of handwritten text lines as transcription source in a crowdsourcing platform [20]. Another use of crowdsourcing in cultural heritage context regards tagging or captioning images. Despite advances in the field of content-based image retrieval, human intake recovers the higher semantic level within cultural heritage image databases, gradually shifting users from passive consumer to pro-active users (manipulating data, improving information retrieval, etc.) [21].

Crowdsourcing is a flexible model that can be applied to a wide range of activities among which adding content to maps like in [22], where it has been developed an application that suggests LPOI (Local Point Of Interest) around the user, through geolocation. Flickr was used to retrieve resources and Wikipedia was used in order to add information. Once the LPOI are chosen, the system computes and displays the suggested path. Every photo retrieved from Flickr is consider as "candidate" until it reached a majority vote, thanks to users' feedback. Then, such a picture becomes eligible of being (or not) displayed whenever a search result encompasses the related LPOI. The purpose of this project is to get users' feedback to improve results displayed by the system.

In the cultural heritage context, the use of digital technologies is increasingly frequent with the aim of enlarging the use of information based on multisensory and multimodal interaction mechanisms and involving user in the exploration of content in a proactive way. An example is 'Youtube Play', a project promoted by Solomon R. Guggenheim, Youtube and HP: some videos selected from a contest were projected from 22 to 24 October 2010 on the facade and in the inner round of the Guggenheim of New York [23]. A similar project was the CITYCLUSTER one [24] that is based on a virtual-reality networking matrix, where interactivity, graphic, and content style coexist in a common virtual territory. Such a project is connected through a high-speed network in which shared environments (both real and imagined) enable remote participants to collaborate, interact, and work together in a common virtual space over distance in real time.

In this context, another concept that can play a key role is gamification [25]. In particular, gamification of micro-tasks spur people coming back for more [26]. Obviously, the main goals of crowdsourcing concern the possibility to create innovative solutions or saving money, but organize required activities more like playing a game can help to remove concerns from people who make tasks for free or for small payment. In fact, gamification has a positive impact on crowdsourcing, both from qualitative and qualitative points of view [26,27].

3 Our Approach

In the context of cultural heritage within a urban area, there are different types of data. On the one hand there are official data coming from GLAM (Galleries, Libraries, Archives, Museums), public administrations, catalogs, and foundations. On the other hand, there are private citizens who, thanks to web 2.0, social networks, and crowdsourcing, have become *prosumer*, as being provider and consumer at the same time, adding contents (photos, tag, feedback, comments) on the Web. Therefore, a large quantity of data that can be found on crowdsourcing platforms, which could support and enrich information already given by official entities. In the context of the SACHER project (Smart Architecture for Cultural Heritage in Emilia Romagna[3], which is co-funded by the Emilia Romagna Region through the POR FESR 2014–2020 fund - European Regional Development Fund), we aim to integrate such two categories of data so as to provide immediate information to users who otherwise would have to do a much longer research. In this paper, we present a first prototype devoted to verify the possibility to gather information from some different cultural heritage sources, with specific regards to social media.

We exploit an open data WebGIS ("Patrimonio culturale dell'Emilia Romagna", that is the Italian for Cultural Heritage of Emilia Romagna region), which lets users visualize a map with points of interest related to cultural heritage (i.e. museums, palazzos, churches, monuments, etc.) and their relative information made available by the Segretariato Regionale for Emilia Romagna region of MiBACT (Minister of cultural heritage and activities and tourism). These open data include geo-localized architectural and archaeological goods. This WebGIS is constantly updated and it is growing: at the time we are writing, it includes 9,092 points of interest.

As crowdsourced data, we take into account data coming from social networks, mainly following two different approaches: *API*, provided directly by the social network considered and *Ad-hoc script*, when official APIs are not provided or if they too limited to be efficiently used. In particular, we aim to select and collect images shared on social media, recording title, description, and hashtags for each one. Our idea is to regularly acquire new images related to some specific case studies, uploaded by users on the different social networks we are monitoring, in order to test the feasibility of scraping those kinds of content from such platforms.

4 Implementation

The architecture of this project is designed bearing in mind the possibility of using both public and open data, like data from public administrations and GLAM and crowd-data from social media, gathering and integrating together to enrich and improve the information available to any users about specific cultural places and pieces of art. The experiment has been designed by considering the following different components, depicted in Fig. 1:

- *Open data:* these data are provided from API offered by different sources like public administrations, GLAM, etc. We stored them into database located in our server.

[3] http://www.eng.sacherproject.com/.

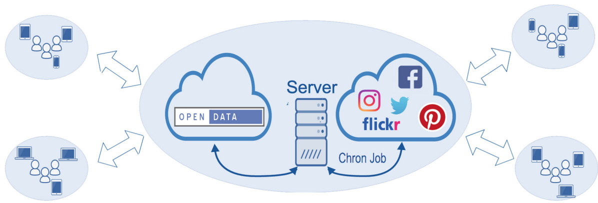


Fig. 1. Software architecture

- *Crowd data:* these data are provided from scraping Python scripts run through a chron-job, keeping continuously updated data from social media, such as Facebook and Twitter.
- *Server:* the server receives data and stores them in a dedicated database. The server is also used in order to reply to users' requests.

In particular, this work focuses on crowd data taken down from five of the most influent and commonly used social networks: Facebook, Twitter, Instagram, Pinterest, and Flickr. In order to develop scraping scripts, Python 2.x and 3.x have been used. Some official APIs (provided by social networks) have been exploited, as well as some specific Python modules, including:

- *Requests v2.18.2*[4], it is a simple module that allows to send HTTP/1.1 requests. It is possible to add parameters, form data, headers, multi-part files with simple Python dictionaries and to access the response data in the same way.
- *Urllib and Urllib2*[5], they are similar to the previous one, but they are a part of the Python Standard Library. Urllib and Urllib2 are modules that allow HTTP/1.1 requests.
- *BeautifulSoup v4.4*[6], it is a Python module for pulling data out of HTML and XML files. BeautifulSoup parses data acquired by Requests or Urllib/Urllib2 modules, which allows searching, navigating, and modifying the parsed tree.
- *Selenium Webdriver v3.4.0*[7], it is a browser automation tool, which automates the control of a browser, so that repetitive tasks can be automated. It has been used together with BeautifulSoup to parse and therefore automate scraping phases.
- *Tkinter*[8], it allows to create GUI in order to manage every scraper scripts.

Each Python script manages to find images and their descriptions as the result of a specific query. In particular, the scripts we have implemented are devoted to collect information about two of the most famous monuments in the city of

[4] https://github.com/requests/requests.
[5] https://docs.python.org/3.6/library/urllib.html#module-urllib.
[6] https://github.com/newvem/beautifulsoup.
[7] https://github.com/SeleniumHQ/selenium.
[8] https://docs.python.org/3.6/library/tkinter.html#module-Tkinter.

Bologna: “Fontana del nettuno” and “Palazzo del podestà”. Furthermore, data were collected at regular intervals in order to investigate how social networks could be exploited as a source of crowdsourced information about cultural heritage.

5 Results and Discussion

In this work, we have taken into account and monitored crowdsourced information about “Fontana del nettuno” and “Palazzo del podestà”, within the following social networks:

- *Flickr:* it was used Flickrapi[9], a native API from Flickr; in this way it was possible to scrape and download a large amount of images, as shown in Figs. 2 and 3;
- *Facebook:* thanks to Facepy[10], a Python library that makes easy to interact with Facebook’s Graph API. We managed to scrape images though in limited manner.
- *Twitter:* TwitterSearch API[11] are used in order to download data from the result of a specific query.
- *Instagram:* it was created an ad hoc Python script that uses Selenium Webdriver library. Selenium takes control of the browser in order to automatize descriptions and images download.
- *Pinterest:* it was implemented an hoc Python script that uses BeautifulSoup and Requests libraries. Pages are first requested, then their DOMs are parsed in order to find all the needed elements.

We scraped images for twelve consecutive days, starting from the 12th of June, until the 24th of June. We searched all the posts/tweets/pins/photos with the hashtags #fontanadelnettuno and #palazzodelpodestà. Results related to the hashtag *fontanadelnettuno* scraped from each monitored social network are reported in Fig. 2, while data shown in Fig. 3 are related to the results obtained for the hashtag *palazzodelpodestà*. In both cases, the social networks with most images are Instagram and Flickr along the monitored period. This is not surprising, since they are platforms mainly devoted to share pictures. The other analyzed social networks reported results pretty far from these two ones.

Some considerations arose from this first experiment and its results:

- We searched items with a particular hashtag and this could be limiting. A first improvement could be searching for the name of the monument (or of the piece of art) also in the title and in the description. Moreover, the geographic coordinates could be considered.

[9] https://www.flickr.com/services/api/.
[10] https://github.com/jgorset/facepy.
[11] https://github.com/ckoepp/TwitterSearch.

– Palazzo del podestà has a final accent on the "a" letter. Users could use hashtags without it. So, for each monument could be searched hashtags with its name put in different forms, for example without the accent or without the articulated prepositions "del". Hence, limits related to languages can occur.
– The last consideration regards the uniqueness of the names of the monuments. In fact, there is also a Fontana del Nettuno in Florence and there are other Palazzo del Podestà in many Italian cities. In these cases a filter on the geographic coordinates could be very useful.

	Flickr	Facebook	Twitter	Instagram	Pinterest
12-Jun	4263	12	244	6861	1250
13-Jun	4263	12	244	6865	1276
14-Jun	4263	12	244	6866	1278
15-Jun	4263	12	244	6868	1279
16-Jun	4263	13	244	6868	1279
17-Jun	4311	13	244	6871	1279
18-Jun	4311	13	244	6876	1290
19-Jun	4315	13	244	6881	1290
20-Jun	4318	13	244	6883	1294
21-Jun	4318	13	244	6886	1294
22-Jun	4332	13	244	6887	1311
23-Jun	4340	13	244	6900	1313
24-Jun	4357	13	244	6902	1315

Fig. 2. Fontana del Nettuno's scraping results: total number of pictures retrieved on a given day in a given data source (Flickr, Facebook, Twitter, Instagram, Pinterest)

	Flickr	Facebook	Twitter	Instagram	Pinterest
12-Jun	373	122	69	877	112
13-Jun	373	122	69	877	112
14-Jun	374	122	69	877	115
15-Jun	374	122	69	877	121
16-Jun	374	122	69	877	121
17-Jun	378	122	70	877	128
18-Jun	378	122	70	877	128
19-Jun	378	122	71	879	128
20-Jun	378	122	71	881	137
21-Jun	378	122	71	882	139
22-Jun	378	122	71	882	139
23-Jun	378	122	71	882	144
24-Jun	380	122	71	883	146

Fig. 3. Palazzo del podestà's scraping results: total number of pictures retrieved on a given day in a given data source (Flickr, Facebook, Twitter, Instagram, Pinterest)

6 Conclusions and Future Works

This first experiment confirms that social networks could be a useful source of information and media in the cultural heritage field, hence the could be used as a crowdsourcing platform and as a data source, to be integrated with other official ones. This work highlights the non-feasibility of the ad hoc scripts in a long term perspective. In fact, a month later the tests, some of the ad hoc scripts fails, because the DOM was changed. Despite this, it is the only possible approach where the APIs are not provided or are too limited. In the first case, however, it's not long-term sustainable. In the second one, instead, an hybrid approach could be use. A first scraper could be made through the use of an ad-hoc script. Then APIs, albeit limited, could be used cyclically to get the new updates since the daily amount of data can be managed.

Interesting future works may include the use of techniques of image recognition with a double purpose. The first one is to understand if the monument is visible or not in the shared pictures (tagged with the related hashtags). The second one is to detect faces in order to crop or hide them for privacy reasons. In the perspective of crowdsourcing, a sentiment analysis could be done on descriptions of the photos, so as to get a raw evaluation about the visitors' perception and feelings about each monument.

References

1. Roccetti, M., Ferretti, S., Palazzi, C.E., Salomoni, P., Furini, M.: Riding the web evolution: from egoism to altruism. In: Proceedings of IEEE Consumer Communications and Networking Conference (CCNC 2008), pp. 1123–1127. IEEE (2008)
2. Toffler, A.: The Third Wave. William Morrow and Company. Inc., New York (1980)
3. Benjamin, W.: The Work of Art in the Age of Mechanical Reproduction of Art. Illuminations, Essays and Reflections (1936)
4. Dusi, N., Ferretti, I., Furini, M.: A transmedia storytelling system to transform recorded film memories into visual history. Entertain. Comput. **21**, 65–75 (2017)
5. Rosner, D., Roccetti, M., Marfia, G.: The digitization of cultural practices. Commun. ACM **57**(6), 82–87 (2014)
6. Mirri, S., Prandi, C., Roccetti, M., Salomoni, P.: Handmade narrations: handling digital narrations on food and gastronomic culture. J. Comput. Cult. Herit. (JOCCH) **10**(4), 20 (2017). Article No. 20
7. Guarnieri, A., Pirotti, F., Vettore, A.: Cultural heritage interactive 3D models on the web: an approach using open source and free software. J. Cult. Herit. **11**(3), 350–353 (2010)
8. Ghini, V., Salomoni, P., Pau, G.: Always-best-served music distribution for nomadic users over heterogeneous networks. IEEE Commun. Mag. **43**(5), 69–74 (2005)
9. Ott, M., Pozzi, F.: Towards a new era for cultural heritage education: discussing the role of ICT. Comput. Hum. Behav. **27**(4), 1365–1371 (2011)
10. Howe, J.: The Rise of Crowdsourcing. su Wired, giugno (2006)
11. Estellés-Arolas, E., González-Ladrón-De-Guevara, F.: Towards an integrated crowdsourcing definition. J. Inf. sci. **38**(2), 189–200 (2012)
12. Heipke, C.: Crowdsourcing geospatial data. ISPRS J. Photogramm. Remote Sens. **65**(6), 550–557 (2010)
13. Mirri, S., Prandi, C., Salomoni, P., Callegati, F., Melis, A., Prandini, M.: A service-oriented approach to crowdsensing for accessible smart mobility scenarios. Mob. Inf. Syst. **2016** (2016)

14. Furini, M., Mandreoli, F., Martoglia, R., Montangero, M.: The use of hashtags in the promotion of art exhibitions. In: Grana, C., Baraldi, L. (eds.) IRCDL 2017. CCIS, vol. 733, pp. 187–198. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68130-6_15
15. Gerla, M., Weng, J.T., Pau, G.: Pics-on-wheels: photo surveillance in the vehicular cloud. In: 2013 International Conference on Computing, Networking and Communications (ICNC), pp. 1123–1127. IEEE (2013)
16. Stathopoulou, E.K., Georgopoulos, A., Panagiotopoulos, G., Kaliampakos, D.: Crowdsourcing lost cultural heritage. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **2**(5), 295 (2015)
17. Kaufmann, N., Schulze, T., Veit, D.: More than fun and money. Worker motivation in crowdsourcing-a study on Mechanical Turk. In: AMCIS, vol. 11, no. 2011, pp. 1–11 (2011)
18. Doan, A., Ramakrishnan, R., Halevy, A.Y.: Crowdsourcing systems on the world-wide web. Commun. ACM **54**(4), 86–96 (2011)
19. Sánchez, J.A., Mühlberger, G., Gatos, B., Schofield, P., Depuydt, K., Davis, R.M., Vidal, E., de Does, J.: tranScriptorium: a European project on handwritten text recognition. In: Proceedings of 2013 ACM Symposium on Document Engineering, pp. 227–228. ACM, September 2013
20. Granell, E., Martínez-Hinarejos, C.D.: A multimodal crowdsourcing framework for transcribing historical handwritten documents. In: Proceedings of 2016 ACM Symposium on Document Engineering, pp. 157–163. ACM, September 2016
21. Van Hooland, S.: From spectator to annotator: possibilities offered by user-generated metadata for digital cultural heritage collections. In: Proceedings of CILIP Conference on Immaculate Catalogues: Taxonomy, Metadata and Resource Discovery in the 21st Century, September 2006
22. Bujari, A., Ciman, M., Gaggi, O., Palazzi, C.E.: Using gamification to discover cultural heritage locations from geo-tagged photos. Pers. Ubiquit. Comput. **21**(2), 235–252 (2017)
23. Gladysheva, D., Verboom, J., Arora, P.: The art tube: strategies, perceptions and outcomes of museums' online video portals. Digit. Cult. Educ. **6**(4), 393–408 (2014)
24. Fischnaller, F., Hill, A.: CITYCLUSTER "from the renaissance to the megabyte networking age": a virtual reality and high-speed networking project. Presence **14**(1), 1–19 (2005)
25. Prandi, C., Nisi, V., Salomoni, P., Nunes, N.J.: From gamification to pervasive game in mapping urban accessibility. In: Proceedings of 11th Biannual Conference on Italian SIGCHI Chapter, pp. 126–129. ACM (2015)
26. Morschheuser, B., Hamari, J., Koivisto, J.: Gamification in crowdsourcing: a review. In: 2016 49th Hawaii International Conference on System Sciences (HICSS), pp. 4375–4384. IEEE, January 2016
27. Prandi, C., Roccetti, M., Salomoni, P., Nisi, V., Nunes, N.J.: Fighting exclusion: a multimedia mobile app with zombies and maps as a medium for civic engagement and design. Multimed. Tools Appl. **76**(4), 4951–4979 (2017)
28. Yuen, M.C., King, I., Leung, K.S.: A survey of crowdsourcing systems. In: 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE 3rd International Conference on Social Computing (SocialCom), pp. 766–773. IEEE, October 2011
29. Brabham, D.C.: Crowdsourcing as a model for problem solving: an introduction and cases. Convergence **14**(1), 75–90 (2008)

Smart Mobility and Sensing: Case Studies Based on a Bike Information Gathering Architecture

Davide Aguiari, Chiara Contoli, Giovanni Delnevo, and Lorenzo Monti(✉)

Interdepartmental Centre for Industrial ICT Research - CIRI ICT,
Università di Bologna, Bologna, Italy
{davide.aguiari2,chiara.contoli,giovanni.delnevo2,
lorenzo.monti20}@unibo.it

Abstract. Mapping services and travel planner applications are experiencing a great success in supporting people while they plan a route or while they move across the city, playing a key role in the smart mobility scenario. Nevertheless, they are based on the same algorithms, on the same elements (in terms of time, distance, means of transports, etc.), providing a limited set of personalization. To fill this gap, we propose PUMA, a Personal Urban Mobility Assistant that aims to let the user add different factors of personalization, such as sustainability, street and personal safety, wellness and health, etc. In this paper we focus on the use of smart bikes (equipped with specific sensors) as means of transports and as a mean to collect data about the urban environment. We describe a cloud based architecture, personas and travel scenario to prove the feasibility of our approach.

Keywords: Smart mobility · Cloud based architecture
Crowdsensing · Crowdsourcing · Personal travel planner

1 Introduction

Smart mobility is playing a strategic role in our daily life in the urban scenario, taking into account that most people live in cities [1,2]. Thanks to the wide diffusion of mobile devices, several services and applications based on the geographical position of users are now available [4], which support citizens while they move across the urban environment [3]. In this context, personalization can be a key factor, enabling independence of citizens, despite some specific conditions (i.e. disabilities) [5] or means of transport [6]. Existing mapping services (e.g. Google Maps) and travel planners (e.g. OpenTripPlanner, Graphhopper) provide multimodal paths computed by the same algorithms, on the basis of the same elements: time, distance, and a limited set of means of transports (e.g. cars, public buses and metros, feet). However, it is not possible to add different factors of personalization, in terms of sustainability, street and personal safety, wellness and health, mood and satisfaction, accessibility. In this context, our idea is to

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 112–121, 2018.
https://doi.org/10.1007/978-3-319-76111-4_12

design a system which acts as a Personal Urban Mobility Assistant (PUMA), supporting citizens in:

- collecting different information about the urban environment (by means of crowdsourcing and crowdsensing activities) in terms of: pollution, traffic, safety, health and fitness, etc.;
- exploiting gathered data, proving multimodal and multipreference paths in the urban environment.

In order to provide multimodal and multipreference paths in urban environments, a detailed mapping of all the elements that affect these factors (e.g., data about pollution, urban barriers and facilities, street lights, data about car accidents, data about crimes, etc.) is needed. Moreover, given out this information, it is necessary a system that lets each user customize and modulate the route computation on the basis of his/her own needs, instead of using the same algorithm for all. As regards the mapping, our approach is based on an open and participatory sensing and mapping system, with low cost sensors, which exploit users' devices too, in a common and shared data repository. We would take advantages by the potentialities of cloud architecture to create a modular open and crowdsourced system. In our work, we also tackled the following challenges:

- To introduce an innovative users' approach towards mobility choices that matches all impact factors for transportation, driving different transportation services, from single to shared, going next to the common existing booking systems, offering a social environment to share experiences and information on sustainable mobility and participate to challenges, info on traffic, lane condition and pollution [10], with the aim of supporting and improving eco-driving and sustainable behaviours [11].
- To develop a smart urban approach to mobility based upon way of booking transport systems that also take into account the carbon footprint [12].
- Integration of sustainable fleets with public transportation (i.e. buses and train) with the possibility to buy tickets by smart payment systems too.
- Data storage and data management (integration from different data sources: from public transportation and route conditions to air pollution obtained by sensors installed on bicycles or other vehicles [13], integration with traffic info).

In this paper, we focus on a specific means of transports: bikes [14]. Bikes can be equipped with different kinds of sensors and can be connected each other, so as to create a specific vehicular network, integrated with the urban infrastructure [7,8] and networks [9] thanks to a cloud architecture. The paper describes the system architecture and details personas and related scenarios, showing how it can be exploited by different users, with different needs and preferences, applying an *altruistic IoT* approach [15].

The remainder of the paper is structured as follows. Section 2 describes the system architecture. Section 3 defines some personas and Sect. 4 presents some use scenarios. Finally, Sect. 5 concludes the paper highlighting some final remarks and future work.

2 Cloud Architecture

In this section, we introduce our system architecture, specifically thought for bicycle vehicles, named BIGA (Bike Information Gathering Architecture), shown in Fig. 1. Our previous work [6] focused mainly on the adoption and implementation of a specific software engineering model that envisioned every component of a mobility application as a service; the reference model was based on microservices, therefore the SMAll architecture was tailored at proposing the definition of an open and standard interface for service access. Instead, BIGA architecture describes at high level the physical and software architecture that might be put in place to provide smart bicycle services. This means that BIGA might be adopted to host and provide the implementation proposed in the SMAll project. Bikes are equipped with sensing devices capable of gathering different kinds of information (data) not only from the environment (e.g., air pollution [16]), but also from the bike itself (e.g., traveled kilometers via odometer). Once (periodically) collected, such information are sent to an entity devoted either to provide connectivity or forwarding data to the cloud via Internet; this entity might be an infrastructure located along the road, such as for example a roadside unit (RSU), or a specific gateway. Cloud infrastructure is where data are processed, stored and made available for being consumed by multiple users.

The idea is to allow users, who are interested in gathering information, to personalize a plan for a given path, depending on their daily habits or needs. The cloud hosts the software that provides path customization and other useful services, but targeted for different uses, as described in the next section. This implies that different applications might require different ways of data collection. Therefore, multiple users can source information by accessing, for example, a web application in order to properly plan in advance their path, or use a mobile application on the smartphone not only for a priori decision, but for real time consultation as well. This means that data can be consumed prior and/or during the journey. At the end of the journey, users can decide to share their experience, i.e., share "collected information" along the path; this would allow to enrich databases with new information, thus resulting in improved bikers experience that can benefit of feedbacks coming from the community. Therefore, with our approach, users are both consumer and producer, so that data are both gathered and disseminated from/in the community.

The vision is that the Municipal District of Bologna (MDB) might act as service provider, i.e., providing to citizenship such "smart bicycles sharing" system. Smart bicycles would be equipped with devices targeted for the different applications (e.g., air pollution monitoring, fitness monitoring, personal safety and carbon footprint). MDB would then rent cloud resources at an infrastructure provider, whose goal would be to provide computational, network and storage resources to have service in place, besides the mobile and web application needed to interact with the service by remote users. Making the cloud hosting the applications, and making these applications available to customers, allow our "Bike as a Service" to fall under the hat of Software as a Service (SaaS). Indeed, cloud approach adoption brings several benefits:

- MDB has no need to install and run applications on their own computers, resulting in less expense in terms of buying new hardware, infrastructure provisioning and consequent maintenance.
- Other emergent paradigms might be put in place on need: Network Function Virtualization (NFV) and Software Defined Networking (SDN) might be adopted in synergy with cloud in order to provide flexible, programmable and cost effective solutions [17]. For example, software applications might run on a Virtual Machines (VMs) interconnected by a virtual network [18]; NFV would help in delivering services as virtual functions, while SDN would help in flexibly managing the (virtual) network [19].
- Cloud approach also calls for service on-demand model, that is, virtual functions could be instantiated, removed or migrated across the network without the need of deploying new hardware.

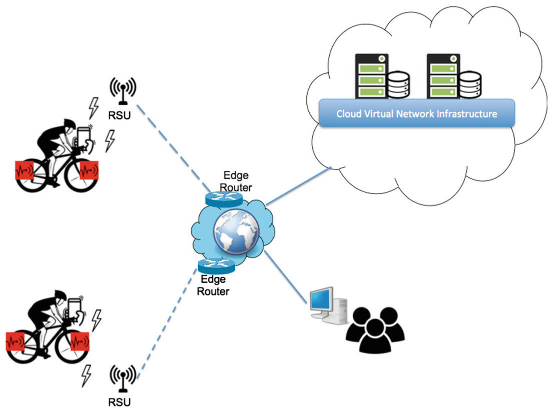


Fig. 1. System architecture

3 Personas

In this Section, we present three personas designed with the aim of defining scenarios exploiting different characteristics of our system.

3.1 Wei

Wei is a Chinese-American visiting Scholar, working to a joint research in Bologna for three months. His research interests are related to climate change and he is very committed in reducing his individual carbon footprint. These days in Bologna are mainly devoted to work and complete all the scheduled research tasks and experiments. In his free time, he likes to go around the city centre and explore the old town of Bologna (Fig. 2).

Wei works at the University of California, San Diego since 2007. He is in Bologna now for a joint research on climate change in the Department of Biological, Geological and Environmental Sciences of the Bologna University. Wei's family is based in S. Diego, where his wife Xiu Ying and his two children (Sean and May) live. They keep in touch with a daily call and Wei sends them a lot of pictures taken wandering around the city while he rides his bike. While in Bologna, Wei lives in the University guest quarters (Residenza di San Giovanni in Monte) located in a prominent monumental complex belonging to the University of Bologna, in the old town centre of Bologna. Wei uses a good trekking bike, loaned by a colleague from the Department for his stay in Bologna. He sporadically uses bus and other public means to reach destination that are too far from the city centre to be reached by bicycle. This responds both to Wei commitment to use sustainable transport and to his travel needs.

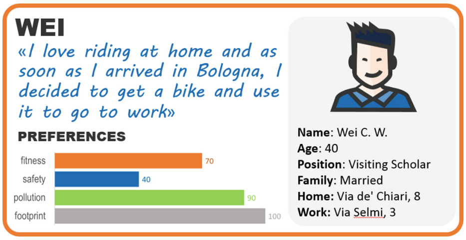


Fig. 2. Wei

3.2 Sven

Sven is a Swedish Erasmus student, living in Bologna for 6 months to complete his master thesis. He is vegan and he is obsessed with fitness. Stay in very good shape, eat vegan and whole food, have a generally healthy life style are very important goals to Sven. During his stay in Bologna, Sven goes to the Department to work to his thesis, to the gym to do his workout, without forgetting to go around with friends, having fun and enjoying the city life as all students do.

He is studying Film Directing at the Faculty of Fine, Applied and Performing Arts at the University of Gothenburg and he is completing his thesis on the Kill Bill movie series by Quentin Tarantino, working in the Department of Arts of the Bologna University. Sven is single. His family of origin lives in Hästevik, a small town near Gothenburg. They keep in touch on a weekly basis with a conference call. While in Bologna, Sven lives in a shared apartment in a neighbourhood outside the city centre. He decided for this location to share the room with his friend Hugo, who is taking his master degree in Economics and Finance in Bologna. The apartment is quite near to the Business and Economics School, where Hugo studies, but pretty far (about 4 km) from the Arts Department. Sven bought a cheap used mountain bike from another Erasmus student leaving Bologna few days after his arrival. He uses a mix of bike and bus to move around the city, depending from weather, time of the day, distance of the destination, but the bicycle is the most used mean to go to the Department and to the Gym on a daily basis because it represents an opportunity to do more workout and also to save money (Fig. 3).

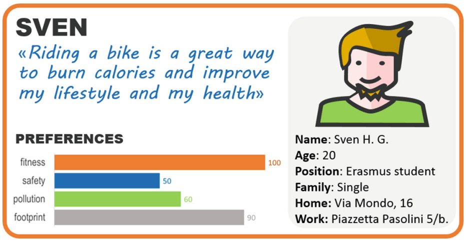


Fig. 3. Sven

3.3 Elena

Elena works full time at the University of Bologna. She was born in Bologna and she grow up in its city centre. Since her husband works in a nearby city, she is in charge of managing their son Tommaso (Tommy) and travel with him to and from school, or to and from her parents' house. While Elena prefers to use the car to reach points of interest outside the centre, when she goes downtown (to work, to Tommy's school and to her parents' house), she prefers to leave the car at home. Elena likes to ride the bike, but she is very worried for Tommy, by the safeness of the travel, weather issues and pollution that can be dangerous, especially during the winter.

Elena works for the University of Bologna since 2008. She works in the International Desk, providing information to students wishing to enroll at the University. She works since 8.30 AM to 4.30 PM for 5 days a week, having a fast lunch in the office nearby. Elena is married with Alberto since 2010 and they have a son, who is 4 years old. Tommy goes to a primary school (Betti Giaccaglia Plesso 2) located in the Montagnola Park, immediately near the Bologna Station. Elena got the option to enroll Tommy in this school because her father and mother live nearby and they are used to pick up Tommy from school every day at 4 PM. Tommy waits for Elena in grandparents' house, located in via Mascarella, for about an hour, to come back home with her. Elena lives in Bolognina, a neighbourhood outside the city centre, quite near to the train station. His husband, Alberto, works in Cesena and this location was chosen mainly to meet his need to easily reach the station. The place is not far from Tommy's School (about 1 km) and from Elena workplace (about 2 km), hence Elena uses a new red city bike, fully equipped with lights and reflectors, to enhance safeness of the travel, and with a baby seat on the back (Fig. 4).

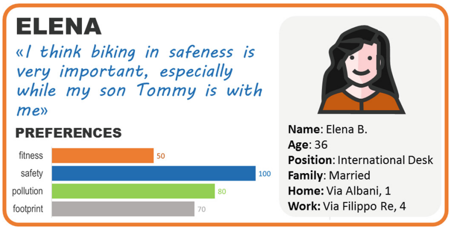


Fig. 4. Elena

4 Travel Scenarios

Travel scenarios related to the personas introduced in the previous section are described in the following subsections.

4.1 Wei

Wei is going to the weekly meeting of the research team, to reschedule some late experiments. It's a foggy day, but despite cold and humid, Wei is happy to take the bike. The meeting will be at 9.00 AM, Wei is leaving the University guest quarters early, so as to have a sweet breakfast in a bar near the Department without being in a hurry. Having more time than what is strictly required to reach his destination, he decided to enjoy the ride and cross the city mainly passing through restricted traffic zones.

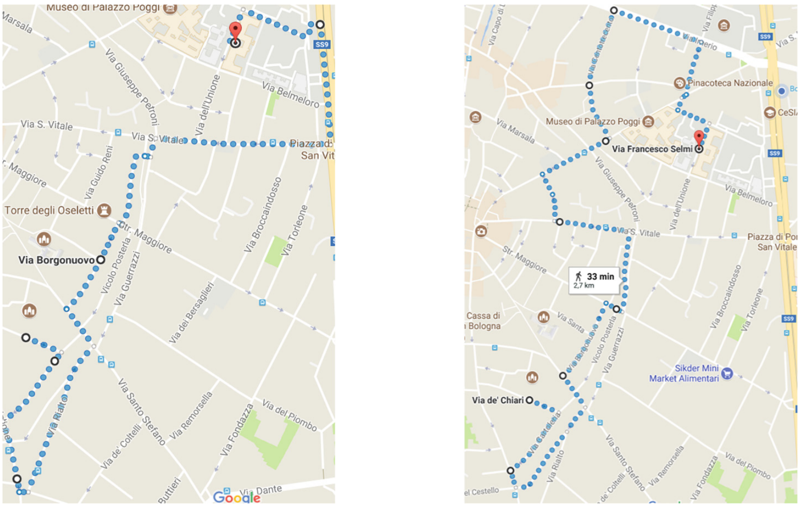


Fig. 5. Wei's routes

In Fig. 5, there are two different paths between Wei's starting point (Via de Chiari) and destination (Via Selmi). On the left side, there is the default and shortest path proposed for bikes by GraphHopper. On the right side, there is the personalized one based on the user's preferences, computed by our PUMA. This latter avoids one of the most congested and polluted roads of Bologna.

4.2 Sven

This morning Sven is going the Cineteca di Bologna to study some sources and will reach at noon his master thesis supervisor at the Department. The weather is not perfect, it is partially cloudy, but Sven prefers to use the bike because he will not have enough time for the Gym. So Sven prefers a longer path so as to do a good workout. Our PUMA proposes a longer path, a path through different green areas and some slopes, as shown in Fig. 6.

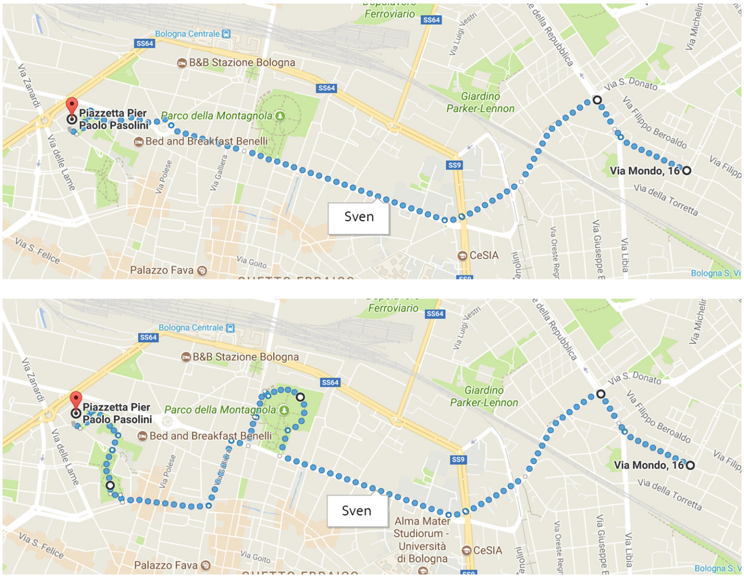


Fig. 6. Sven's routes

4.3 Elena

Elena is going to work, she will leave Tommy at the school on her way. The weather is very good, 25 °C, a perfect spring day with a perfect temperature. She is leaving in at 7.45 AM, just in time to stop at the School, say bye to Tommy and go to work in schedule without being in a hurry. She decides to go safe using the available bicycle lanes and to select a route through parks and green areas to enjoy the spring weather and avoid a large exposure to pollutant. Our system proposes a path that does not cover Via Irnerio (as shown in Fig. 7), a route not safe for cyclist because of the traffic, that includes cars and buses, and not clean, due to the pollution produced by these means of transport.

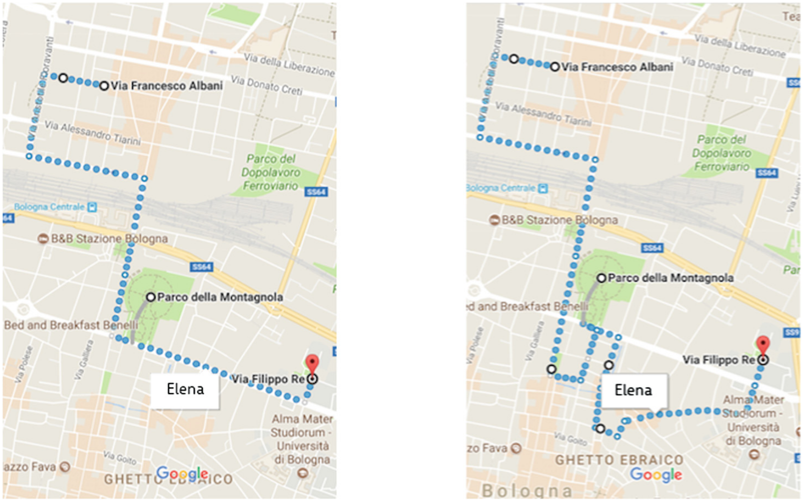


Fig. 7. Elena's routes

5 Conclusions

This paper presents a Personal Urban Mobility Assistant (PUMA) system, based on the idea of letting citizens and public administrations gathering and exploiting integrated information about the urban environment, by using web applications and mobile devices. Users are at the same time producers and consumers of data and services, thanks to a cloud architecture approach, providing typical SaaS benefits. The paper focuses on a specific mean of transport: bikes. It describes three different personas and related scenarios, with the aim of illustrating how our system can support smart and sustainable mobility in a urban scenario, thanks to crowdsourcing and crowdsensing activities.

References

1. Mirri, S., Prandi, C., Salomoni, P., Callegati, F., Melis, A., Prandini, M.: A service-oriented approach to crowdsensing for accessible smart mobility scenarios. Mob. Inf. Syst. **2016** (2016)
2. Melis, A., Mirri, S., Prandi, C., Prandini, M., Salomoni, P., Callegati, F.: Integrating personalized and accessible itineraries in MaaS ecosystems through microservices. Mob. Netw. Appl. 1–10 (2017)
3. Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J.R., Mellouli, S., Nahon, K., Scholl, H.J.: Understanding smart cities: an integrative framework. In: 2012 45th Hawaii International Conference on System Science (HICSS), pp. 2289–2297. IEEE, January 2012
4. Dusi, N., Ferretti, I., Furini, M.: PlayTheCityRE: a visual storytelling system that transforms recorded film memories into visual history. In: 2016 IEEE Symposium on Computers and Communication (ISCC), pp. 85–90. IEEE, June 2016

5. Mirri, S., Prandi, C., Salomoni, P.: Personalizing pedestrian accessible way-finding with mPASS. In: 2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC), pp. 1119–1124. IEEE (2016)
6. Mirri, S., Prandini, M., Salomoni, P., Delnevo, G., Callegati, F., Melis, A.: I want to ride my bicycle: a microservice-based use case for a MaaS architecture. In: 2017 IEEE Symposium on Computers and Communication (ISCC). IEEE (2017)
7. Marfia, G., Pau, G., Giordano, E., De Sena, E., Gerla, M.: Vanet: on mobility scenarios and urban infrastructure. A case study. In: 2007 Mobile Networking for Vehicular Environments, pp. 31–36. IEEE (2007)
8. Roccetti, M., Gerla, M., Palazzi, C.E., Ferretti, S., Pau, G.: First responders' crystal ball: how to scry the emergency from a remote vehicle. In: IEEE International Performance, Computing, and Communications Conference, IPCCC 2007, pp. 556–561. IEEE, April 2007
9. Ghini, V., Salomoni, P., Pau, G.: Always-best-served music distribution for nomadic users over heterogeneous networks. IEEE Commun. Mag. **43**(5), 69–74 (2005)
10. Tse, R., Zhang, L.F., Lei, P., Pau, G.: Social network based crowd sensing for intelligent transportation and climate applications. Mob. Netw. Appl. 1–7 (2017)
11. Di Lena, P., Mirri, S., Prandi, C., Salomoni, P., Delnevo, G.: In-vehicle human machine interface: an approach to enhance eco-driving behaviors. In: Proceedings of 2017 ACM Workshop on Interacting with Smart Objects, pp. 7–12. ACM (2017)
12. Bigazzi, A.Y., Figliozzi, M.A.: Roadway determinants of bicyclist exposure to volatile organic compounds and carbon monoxide. Transp. Res. Part D **41**, 13–23 (2015)
13. Alvear, O., Zamora, W., Calafate, C.T., Cano, J.C., Manzoni, P.: EcoSensor: monitoring environmental pollution using mobile sensors. In: 2016 IEEE 17th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), pp. 1–6. IEEE, June 2016
14. Lerner, W., Van Audenhove, F.J.: The future of urban mobility: towards networked, multimodal cities in 2050. Publ. Transp. Int.-Engl. Ed. **61**(2), 14 (2012)
15. Mirri, S., Prandi, C., Salomoni, P., Monti, L.: Social location awareness: a prototype of altruistic IoT. In: 2016 8th IFIP International Conference on New Technologies, Mobility and Security (NTMS), pp. 1–5. IEEE, November 2016
16. Tse, R., Pau, G.: Enabling street-level pollution and exposure measures: a human-centric approach. In: Proceedings of 6th ACM International Workshop on Pervasive Wireless Healthcare, pp. 1–4. ACM (2016)
17. Callegati, F., Cerroni, W., Contoli, C., Santandrea, G.: SDN controller design for dynamic chaining of virtual network functions. In: 2015 4th European Workshop on Software Defined Networks (EWSDN), pp. 25–30. IEEE (2015)
18. Callegati, F., Cerroni, W., Contoli, C., Santandrea, G.: Performance of multi-tenant virtual networks in OpenStack-based cloud infrastructures. In: 2014 Globecom Workshops (GC Wkshps), pp. 81–85. IEEE (2014)
19. Callegati, F., Cerroni, W., Contoli, C., Santandrea, G.: Implementing dynamic chaining of virtual network functions in OpenStack platform. In: 2015 17th International Conference on Transparent Optical Networks (ICTON), pp. 1–4. IEEE (2015)

3D Interaction with Mouse-Keyboard, Gamepad and Leap Motion: A Comparative Study

Fabio Pittarello(✉), Alexandru Dumitriu, and Elisa Piazza

Università Ca' Foscari Venezia, Mestre-Venezia, Italy
pitt@unive.it, {836110,838565}@stud.unive.it

Abstract. Serious gaming can represent a key for fostering learning and letting children to acquire new information and skills while doing engaging activities. Among the different types of games, those ones based on interactive 3D environments are widely diffused and appreciated. A key component of the design of these experiences is the choice of the input device that will be used by the players and the mapping of the users' intentions to the actions in the 3D environment. The choice of the proper device can lead to benefits in terms of user engagement, which often is the prerequisite for learning. There are also additional dimensions to consider, as the usability and the physical fatigue. Their undervaluation, in an educational context, can hamper the successful outcome of the experience. For this reason in this work we compared the use of three different input devices (a mouse-keyboard set, a gamepad and the Leap Motion, a sensor for recognizing hand gestures) for controlling a 3D educational gaming experience focused on environmental sustainability. We organized a comparative study with 30 children of the Primary School, evaluating the interaction in terms of usability, engagement and physical fatigue. The results evidenced the potential of the Leap for engaging the children, but also drawbacks in terms of usability and physical fatigue that should be taken into consideration for the development of this technology and the design of experiences based on it.

Keywords: 3D interaction · Children · Comparative study
Engagement · Gesture-based input · Physical fatigue · Serious gaming
Usability

1 Introduction

Serious gaming represents an opportunity for permitting the users to acquire new knowledge and skills while performing engaging activities. The creation of an engaging gaming experience passes through different important factors, among which the definition of how the user interacts with the game. The choice of a specific input device is an important choice that can make the difference. While at the beginning the properties of the input devices and their influence on the user experience were evaluated only in terms of usability, as a consequence of

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 122–131, 2018.
https://doi.org/10.1007/978-3-319-76111-4_13

the Taylorist tradition of time and motion studies related to structured and paid work [4], the shift to other human activities determined a growing importance for other parameters, such as the engagement. According to O'Brien and Toms [13] engagement can be defined as the sum of six dimensions: perceived usability, felt involvement, focused attention, novelty, endurability and aesthetics. Please note that this analytical definition includes the usability as well, which accounts for a number of issues, among which the perceived effort in using the application, the ability to accomplish the task and the feeling of being in control [13].

The physical fatigue is another important factor for all those experiences that go beyond desktop computing or that however require an heavy physical involvement. The goal of this paper is to investigate the use of different categories of input devices and related interaction styles in the context of a 3D gaming experience, evaluating their points of strength and weakness in relation to all the parameters cited above. Although this kind of comparative study is not novel [1,10], we introduce a specific focus on mid-air hand gestures and the Leap Motion, a low cost device that permits to track accurately them. Mid-air gestures represent a specific category of gestures that can be used as an input medium and that typically are associated to actions such as confirmation, selection, navigation and modification. This study focuses in particular on mid-air hand gestures which don't take advantage of an assisted physical medium, such as a pen or a remote control [6]. We tried to give an answer to the following research question: how the Leap compares to the mouse-keyboard set and the gamepad for completing common interaction tasks in a 3D environment, in relation to usability, engagement and physical fatigue? The educational game that we designed for comparing these devices was focused on environmental awareness and tested with 30 children of the Primary School aged 10. The study gave mixed results, showing that gesture-based interaction was appreciated for a number of parameters that define the engagement, thus confirming its educational potential. However a number of drawbacks emerged, in terms of usability and physical fatigue, that should be taken in consideration when designing experiences based on this input modality.

2 Related Work

Gesture-based interaction has become popular in the last decade thanks to the proposal of commercial products, targeted mainly to the gaming domain. These devices are usually characterized by a low cost and an increasing level of accuracy and have become a viable alternative to the long-standing and often costly data gloves [3]. The interest for their potential is constantly increasing for different application domains. We can make a rough classification distinguishing systems which require to grab a device that embeds sensors or markers for monitoring the user gestures, such as the Nintendo Wiimote, and systems that rely on an external sensor for tracking the user motion, such as the Microsoft Kinect and the Leap Motion [7]. Many applications of gesture-based technology are related to interactive 3D environments. This is not surprising, given that 3D environments mimic the real world and gesture-based interaction seems an interesting

Table 1. Mapping actions in the 3D world to input devices

Action	Mouse & K.board	Gamepad	Leap motion
Walk forward	Up arrow	Left stick f.ward	Closed hand f.ward
Walk backward	Down arrow	Left stick b.ward	Closed hand b.ward
Turn left	Left arrow	Left stick to the l.	Closed hand to the l.
Turn right	Right arrow	Left stick to the r.	Closed hand to the r.
Grab object	Left mouse butt.	A button	Grab
Ungrab object	Left mouse butt.	A button	Open hand
Raise hand	f.ward mouse wheel	LB button	Raise hand
Lower hand	b.ward mouse wheel	RB button	Lower hand
Move hand forward	Mouse f.ward	Right stick f.ward	Hand f.ward
Move hand backward	Mouse b.ward	Right stick b.ward	Hand b.ward
Move hand to the left	Mouse to the left	Right st. to the l.	Hand to the left
Move hand to the right	Mouse to the right	Right st. to the r.	Hand to the right

opportunity to extend the mimesis also to control it. The design of interaction for 3D worlds requires to manage a number of actions, from the navigation to the manipulation of objects that are contained inside of them. In spite of all the research work done for trying to improve the usability of interactive worlds [11,14], the complexity and the variety of the issues has prevented from coming to a satisfactory solution, available for all the situations. Lapointe et al. [10] in 2011 stated that *research shows that, in the case of 3D interface, there is still not an input device that demonstrates its superiority for accomplishing basic 3D tasks such as navigation, manipulation and selection*. This statement is still true, as shown also by additional surveys [8,9] suggesting that different solutions are suited to different contexts and application domains. In this scenario there are interesting studies that compare the use of different input devices and mappings [1,10]. The study described in this paper belongs to this category and tries to enlarge the comparative approach to other categories of devices not previously considered, such as the Leap, for highlighting points of strength and weaknesses in a specific context of use.

3 Designing the 3D Experience

The educational experience was focused on the exploration of a park (see Fig. 1 on the left). The visitors had to retrieve different types of waste in the shortest time and put them in the proper recycling bin, in order to improve their environmental awareness. Any error performed during the interaction, such as the involuntary ungrabbing of the waste or its association to the wrong recycling bin caused the object to return to its original position.

For giving generality to the study, we established that the user experience should have been based on a set of actions that are usually available for 3D

Fig. 1. The 3D interface and one of the tutorials related to the mouse-keyboard set

games. The first column of Table 1 displays all the available actions. For avoiding to add unnecessary degrees of complexity to the experience, we limited the degrees of freedom for locomotion, considering only forward and backward walking and the possibility to turn left or right. Concerning the interaction with the objects, we considered the actions for controlling a 3D counterpart of the user's hand and grabbing 3D objects. All these actions were mapped to the input devices selected for this study (mouse-keyboard, gamepad, Leap Motion), for granting the same level of expressivity for each situation and easing the comparisons. For mapping the actions to the input devices we took into account the configurations that can be found in gaming, although in some cases we had to choose among possible alternatives. In particular, for the mouse-keyboard, we mapped the locomotion to the four directional arrows of the keyboard, as an alternative to the more common WASD solution, because we wanted to avoid an excessive bias due to previous gaming experiences. For the Leap we couldn't count on any established praxis. Additional restrictions came from the limitedness of its library of gestures. In this case, following the results of studies [2,5] which underline a preference by the users for gestures that mimic the action of the real world, we used *pantomimic* gestures for all the manipulation gestures, including the hand movements and the act of grabbing an object. Where it was not possible, such as for the locomotion, we preferred *deictic* gestures to *symbolic* gestures that have to be learned and whose meaning can vary for different users, contexts and cultures. We chose also to maintain simplicity, avoiding bimanual gestures that would have added a level of complexity. For this reason we designed locomotion gestures that took advantage of the hand already in use for grabbing the objects, for guiding navigation as well.

The output interface, displayed in Fig. 1, was characterized by the subjective view of the user in the 3D world and included a simple HUD showing information such as the score, the number of objects available in the park and collected, the name of the object currently grasped. Textual messages appeared in the center of the screen for underlying important actions, such as the act of placing an object in the proper/improper waste bin and for warning the user if she inadvertently went away from the operation zones. For the manipulation of the objects, the feedback was given by the 3D counterpart of the hand of the user and by the change of color of the object to collect, when the hand had reached the right position for grasping it. The interface was multimodal in that it was complemented by audio tones associated to the main events. Ambient audio effects were added as well, for improving the sense of presence in the 3D world. The system was developed taking advantage of Unity3D, a well-known SDK for gaming development. During the development of the system, a keen attention was devoted to obtain a smooth and precise interaction with all the devices. We dedicated a complementary attention for defining comparable locomotion and manipulation speeds for each device, in order to obtain a fair comparison of the time necessary to complete the experience. During this process we organized also an informal pilot study with two children aged 10 that tried all the different interfaces. Their feedback was very useful for refining the results. At the end of the development we had three distinct and comparable systems with the same functionalities that differed only for secondary issues (e.g., the shape of the 3D hand in the 3D environment).

4 The Comparative Study

We organized the comparative study with the collaboration of the teachers of two classes of the Primary School “S. Giovanni Bosco”, located in a small center of Northern Italy, and the participation of 30 children (13 boys and 17 girls) aged 10. The parents of the children were informed about the goals and the structure of the study and signed an informed consent form before its start.

The experimental setup was based on a laptop connected to an external big screen and to the input devices. Aside from the initial briefing where each child was given an explanation of the goal of the game and of the experiment, the core of the experience were the sessions with the three input devices. We used a within-subjects design, considering the device as the main independent variable. For all these sessions we used the same 3D world, with the objects and the bins placed in the same place, in order to avoid differences related to the difficulty of retrieving or manipulating them. For counterbalancing the learning effect, we divided the 30 children in 6 groups, using for each group a sequence of sessions derived by one of the possibile permutations of the three input devices, as suggested in [12]. An additional care for diminishing the learning effect was the introduction of a practicing trial, performed before the main sessions, for all the devices. Each trial was introduced by a short video tutorial. The miniatures on the right part of Fig. 1 are taken from the video related to the mouse-keyboard set and display the relation between the manipulation of the input device and a number of actions in the 3D world (i.e. go forward, turn left, collect and release the waste). We dedicated about 1 h and half to test each child. We took advantage of questionnaires for collecting qualitative data related to usability, engagement and physical fatigue. Quantitative data, among which the time for completing the sessions and the type and number of errors, were collected through direct observation and video recording. We asked the children fo fill in a preliminary questionnaire before the initial briefing, for collecting demographical data and information about their prior experience related to gaming and input devices. Then we asked the children to fill in an intermediate questionnaire after the session with each device. This questionnaire was organized as a set of closed questions focused on the 6 parameters of the engagement and on the physical fatigue. We formulated the questions related to each parameter in plain terms, suitable to che children's age, and we checked them with the teachers before the study. The children where asked to select among 5 values ranging from *Not at all* to *Extremely*. The answers were then converted in numerical values from 1 to 5 for the following analysis. At the end of the three sessions we asked the children to fill in a final questionnaire that included a set of open questions, for giving them the opportunity to highlight the positive and negative features of the experience and to propose modifications to the mappings as well. Both the user interaction and the output interface were recorded with a digital camera and a video capture software (see Fig. 2). The two streams were then post-processed for obtaining a single synchronized video, useful for re-examining in detail the sessions.

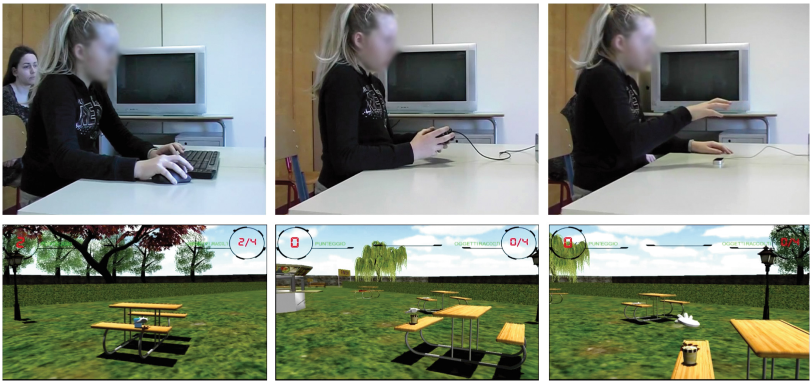


Fig. 2. Interacting with the mouse-keyboard, the gamepad and the Leap

4.1 Results

The exam of the initial questionnaire showed that only 1 child out of 30 didn't own a personal computer, a game console or a multitouch device. Coming to devices for capturing mid-air gestures, only 4 children declared to use them (i.e. 1 Kinect, 3 Wiimotes). Figure 3 resumes the results of the intermediate questionnaires related to the engagement and the physical fatigue. Please note that the question related to the usability was split into 4 sub-questions for getting better insights about the different types of actions. The box plot, which has been extensively used in this paper for visualizing the distribution of the scores (based on a 5 points scale), displays the limits of the first and the third quartile. A thick horizontal line inside the box and a dashed line are used for visualizing respectively the median and the mean. The single dots evidence the outliers

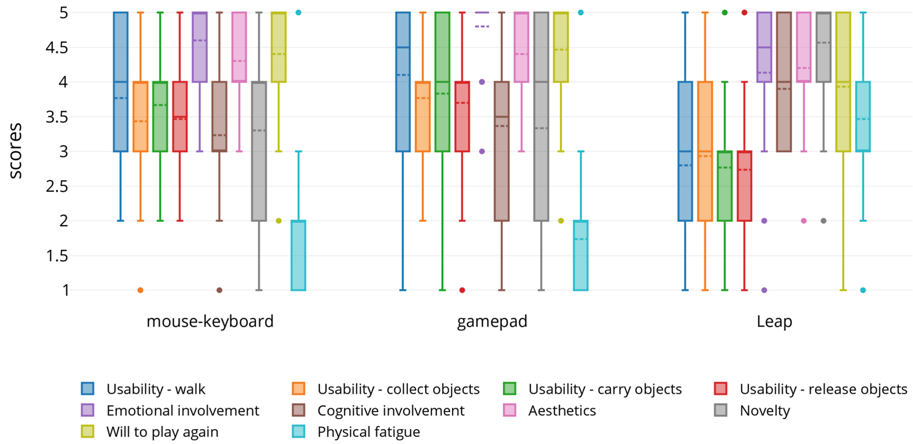


Fig. 3. Results of the intermediate questionnaire related to the engagement and the physical fatigue

as well. Figure 3 shows that gesture-based interaction gained the lowest scores for what concerned the usability, and that some actions were more difficult to perform. It was clearly considered by the children also as the most physically demanding and required also additional cognitive involvement. On the other side, gesture-based interaction was perceived as the most novel. For all the other parameters differences are less evident. The gamepad and the mouse-keyboard obtained generally similar scores, with a slight advantage of the gamepad for all the parameters.

The answers to the open questions brought additional insights. The mouse-keyboard set was appreciated for the familiarity of the devices, although for some children the same familiarity appeared as boring. This input solution received many appreciations for its usability as well, although some children were confused by the use of the keys and suggested to simplify interaction using only the mouse. Some children suggested also to use the more common WASD set of keys for locomotion, as an alternative to the use of the arrows. Also the gamepad was appreciated for its familiarity and usability. A lot of children stated that the gamepad was very engaging. While many children appreciated how we mapped the gamepad buttons and sticks to the actions in the 3D world, some of them suggested alternatives for locomotion (e.g., arrows instead of the stick) and for the hand motion (e.g., using different keys or a stick for changing the hand's vertical position). The Leap was appreciated by the users for its novelty and the possibility to use their hands for controlling the interaction, giving a feeling of direct connection with the 3D world. Different children however emphasized the physical fatigue, due to the need of keeping the arm in a straight position for long periods of time. Children evidenced also the difficulty to interact and perceived a lack of precision or an excessive sensitiveness during the use of the device. On the other side, children were engaged by the challenge of controlling it properly.

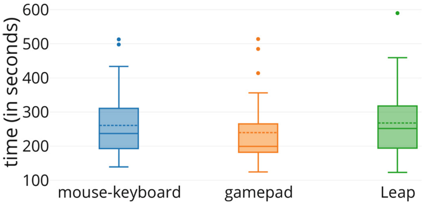


Fig. 4. Time of completion

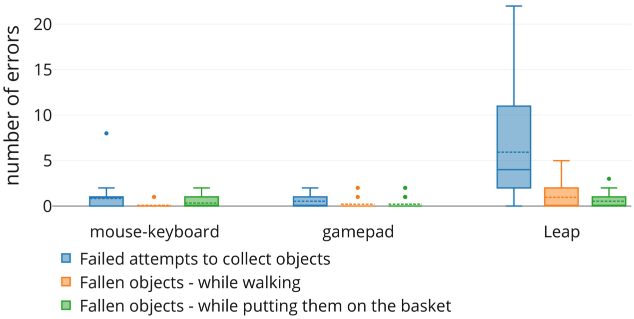


Fig. 5. Errors

The children's suggestions were mainly related to the locomotion actions: some of them suggested to associate them to the act of pointing a finger. One of the children considered the gesture of moving the fist as a combat action and stated that he would have preferred an alternative mapping. Another child suggested to mimic the legs' movement with the fingers.

The analysis of the time needed for completing the session (Fig. 4) shows that the gamepad permitted to obtain the best results in terms of mean, median and quartiles, while the use of the keyboard/mouse and of the Leap obtained similar but worse results. The analysis of errors shows that a much higher number of errors characterized the experience with the Leap (see Fig. 5). The most evident differences are related to the act of collecting the objects. The exam of the errors during the transportation and the placement of the collected object confirms a difficulty in using the Leap, although differences are less evident. Finally, while the mouse-keyboard set and the gamepad required more assistance during the practicing phase (there were respectively 20 and 23 requests of assistance, vs. 16 requests for the Leap), things radically changed during the main sessions. There was only one request of assistance for the mouse-keyboard set and the gamepad, while there were 9 requests for the Leap.

5 Discussion and Conclusion

The result of this study showed that the Leap and gesture-based interaction was perceived as an interesting novelty (Fig. 3), in a context characterized, as shown by the first questionnaire, by the familiarity with gaming platforms and desktop environments, but scarce acquaintance with this type of device and interaction. Children were engaged by gesture-based interaction and by those gestures that mimicked the manipulation in the real world [2,5], as evidenced from the answers to the open questions. All the requests of modifying the mappings of the Leap were referred to the locomotion, which unfortunately was not possible to mimic. A number of issues emerged from this study, related in particular to the physical fatigue and the usability (Fig. 3). Both issues were evidenced also by the answers to the open questions. Besides, the quantitative data suggest that what contributed to originate the children's judgement, rather than an increase in the time needed for completing the task, was the high number of errors performed during the interaction, due to the lack of precision evidenced in the open questions. The direct observation evidenced also that a part of the errors (e.g., the involuntary ungrabbing of the object during locomotion) were due to the fact that the children moved their hand outside the scanning boundaries of the device. The analysis of the assistance given to children evidenced also that the difficulties were due, rather than to a conceptual comprehension of the interaction mechanism, to its practical utilization.

Resuming, in this work we identified some major problems in gesture-based interaction whose resolution could led to a wider use of this technology, that it is interesting for a number of factors. We can infer from the results of the study that the issues related to physical fatigue seems to be related to the nature of gesture-based interaction, while the usability problems seem more related to technological difficulties. Solutions for the first issue can be related to the design of gestures less physically demanding and that don't require a continuous action by the user. These solutions however should take into account the preference for gestures that mimic reality, as confirmed also from this study. In this respect, additional benefits might come from considering, during the design, the posture of the whole body during interaction, in order to define a more satisfying solution from the ergonomic point of view. The availability of a library with a wider number of gestures of course would bring great benefits for further experimentation. Solutions for the second issue can come from different factors, including the tuning of the mapping used in this study, but also the availability of more advanced libraries for the precise management of the input. The results of this study suggest also that there are a number of potential domains for the application of gesture-based interaction, where the impact of physical fatigue would be minor, because the interaction with the system would be limited or more distributed in time. Smart home applications seem good candidates on this respect. Additional insights might come from long term studies, for eliminating completely the novelty effect and give the user more time for improving the control of the device. Finally, concerning the educational potential of this new technology, the study gave mixed results. Only a part of the factors that define the engagement received higher scores (novelty). For other parameters the Leap didn't gave a competitive advantage over the other devices. Some results, such as the cognitive involvement requested by the Leap, might be read as positive in an educational context. As a matter of fact, part of the children declared that they appreciated the challenge of mastering this device. However the prolonged use of the device and the precision problems might hamper the use of gesture-based interaction and discourage the users after the end of the novelty effect. Therefore usability and physical fatigue are critical factors that should be taken in serious account in the future development of this technology. At the current state gesture-based interaction seems to be a great solution for educational experiences that don't require a prolonged use and high levels of precision; in this sense, it appears as a useful integration rather than a replacement of the existing technologies.

References

1. Ardito, C., Buono, P., Costabile, M.F., Lanzilotti, R., Simeone, A.L.: Comparing low cost input devices for interacting with 3D virtual environments. In: 2nd Conference on Human System Interactions (HSI 2009), pp. 292–297. IEEE (2009)
2. Ardito, C., Costabile, M.F., Jetter, H.-C.: Gestures that people can understand and use. J. Vis. Lang. Comput. **25**(5), 572–576 (2014)
3. Dipietro, L., Sabatini, A., Dario, P.: A survey of glove-based systems and their applications. IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev. **38**(4), 461–482 (2008)
4. Dix, A., Finlay, J.E., Abowd, G.D., Beale, R.: Human-Computer Interaction, 3rd edn. Prentice-Hall Inc., Upper Saddle River (2003)
5. Grandhi, S.A., Joue, G., Mittelberg, I.: Understanding naturalness and intuitiveness in gesture production: insights for touchless gestural interfaces. In: CHI 2011, pp. 821–824. ACM, New York (2011)
6. Groenewald, C., Anslow, C., Islam, J., Rooney, C., Passmore, P., Wong, W.: Understanding 3D mid-air hand gestures with interactive surfaces and displays: a systematic literature review. In: 30th International BCS HCI Conference, pp. 43:1–43:13. BCS Learning & Development Ltd, Swindon (2016)
7. Hodson, H.: Leap motion hacks show potential of new gesture tech. New Sci. **218**(2911), 21 (2013)
8. Hoppe, A.H., van de Camp, F., Stiefelhagen, R.: Interaction with three dimensional objects on diverse input and output devices: a survey. In: Stephanidis, C. (ed.) HCI 2017. CCIS, vol. 713, pp. 130–139. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58750-9_18
9. Jankowski, J., Hachet, M.: A survey of interaction techniques for interactive 3D environments. In: Sbert, M., Szirmay-Kalos, L. (eds.) Eurographics 2013 - State of the Art Reports. The Eurographics Association (2013)
10. Lapointe, J.-F., Savard, P., Vinson, N.: A comparative study of four input devices for desktop virtual walkthroughs. Comput. Hum. Behav. **27**(6), 2186–2191 (2011)
11. Lindeman, R.W.: Making VR more usable: the state of effectiveness in virtual reality. J. Vir. Reality Soc. Japan **11**(1), 24–27 (2006)
12. Markopoulos, P., Read, J.C., MacFarlane, S., Hoysniemi, J.: Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers. Morgan Kaufmann Publishers Inc., San Francisco (2008)
13. O'Brien, H.L., Toms, E.G.: The development and evaluation of a survey to measure user engagement. J. Am. Soc. Inf. Sci. Technol. **61**(1), 50–69 (2010)
14. Sebok, A., Nystad, E., Helgar, S.: Navigation in desktop virtual environments: an evaluation and recommendations for supporting usability. Vir. Reality **8**(1), 26–40 (2004)

Discovering the City: Crowdsourcing and Personalized Urban Paths Across Cultural Heritage

Giovanni Delnevo[3], Andrea Melis[1(✉)], Silvia Mirri[2], Lorenzo Monti[3], and Marco Prandini[2]

[1] Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi", Universitá di Bologna, Bologna, Italy
a.melis@unibo.it

[2] Department of Computer Science and Engineering, Universitá di Bologna, Bologna, Italy

[3] Interdepartmental Centre for Industrial ICT Research - CIRI ICT, Universitá di Bologna, Bologna, Italy

Abstract. Travel planners and mobile applications related to cultural heritage can play an interesting role in the development of smart cities, when they are integrated each other, engaging the user in touristic and entertainment activities, letting him/her be a source of cultural resources. This paper focuses on a microservices-based architecture, defined with the aim of providing support in computing personalized urban paths across cultural heritage places and in sharing multimedia resources about points of interest. A prototype of mobile application has been implemented on the basis of such architecture, showing the feasibility of the proposed approach thanks to personas and related scenarios.

Keywords: Smart city · Cultural heritage · Crowdsourcing
Microservices · Personal travel planner

1 Introduction

The wide diffusion of smart objects (including smartphones, tablets, smart watches) is profoundly affecting our daily life, changing our habits and the way we conduct all our activities, from communicating to learning, from working to spending free time, and so on. In this context, urban infrastructures, connectivity, and Internet of Things are supporting and enhancing the application and the achievement of smart city paradigms [1]. As a consequence, we are witnessing the proliferation of several mobile applications devoted to support the users in all those activities in the smart city context. In such a scenario, a key role is played by mapping services (e.g. Google Maps) and travel planners (e.g. OpenTripPlanner, Graphhopper), which support users in exploiting urban paths computed by similar algorithms, on the basis of similar characteristics (i.e. time,

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 132–141, 2018.
https://doi.org/10.1007/978-3-319-76111-4_14

distance, means of transport, etc.). Generally, such kind of applications shows some lacks in terms of personalization, in fact, the user cannot set any preferences in terms of health and fitness issues, personal and street safety, air quality (pollution, dust and pollen, etc.), user interface and interaction (which can be strongly affected by the use of context [2]), and so on.

Another interesting topic which can get great benefits from exploiting mobile applications on smart devices and from user's preferences mechanisms is the cultural heritage one, related to touristic activities. In particular, geolocalization sensors can play a strategic role in this context, enhancing available functionalities and services [3]. Moreover, in the cultural heritage context, information collected from tourists and citizens [4,5] by means of their mobile devices [6], also thanks to their activities on social networks, could be integrated with the data coming from the official sources (such as data coming from municipalities, touristic offices, museums, public administrations and institutions, private foundations, etc.), enriching the whole amount of available resources and knowledge [7,8]. Again, personalization on the basis of users' preferences could offer great benefits in this field.

This work aims to integrate these two realms, by defining a system architecture based on microservices [9]. In particular, in this paper we focus our attention on how to define and to manage users' preferences, detailing the Preferences Layer of a microservices-based architecture. On the basis of such an architecture, we have designed and developed a prototype of mobile application (named Cicerone) devoted to compute personalized urban paths, so as to let the user move across preferred cultural heritage places and Points Of Interest (POIs) in the city. Such an application has been thought to let the user enjoy multimedia resources about the points of interest s/he is visiting, providing additional data (i.e. pictures, comments, reviews), enriching the information available for those POIs, thanks to crowdsourcing activities and to gamification mechanisms. This work has been done within the SACHER project (Smart Architecture for Cultural Heritage in Emilia Romagna[1], which is co-funded by the Emilia-Romagna Region through the POR FESR 2014-2020 fund - European Regional Development Fund). Hence, we have based our prototype on official data sources about cultural heritage in the city of Bologna and in the Emilia-Romagna Region.

The remainder of the paper is structured as follows. Section 2 provides a description of the Preferences Layer Architecture. Section 3 presents a prototype of mobile application (named Cicerone), implemented on the basis of such an architecture, describing two specific personas and related usage scenarios. Finally, Sect. 4 concludes the paper highlighting some final remarks and future work.

2 The Preferences Layer Architecture

Figure 1 shows the proposed preferences layer architecture. The microservice paradigm is at the basis of the design of our architecture, letting it work on

[1] http://www.eng.sacherproject.com/.

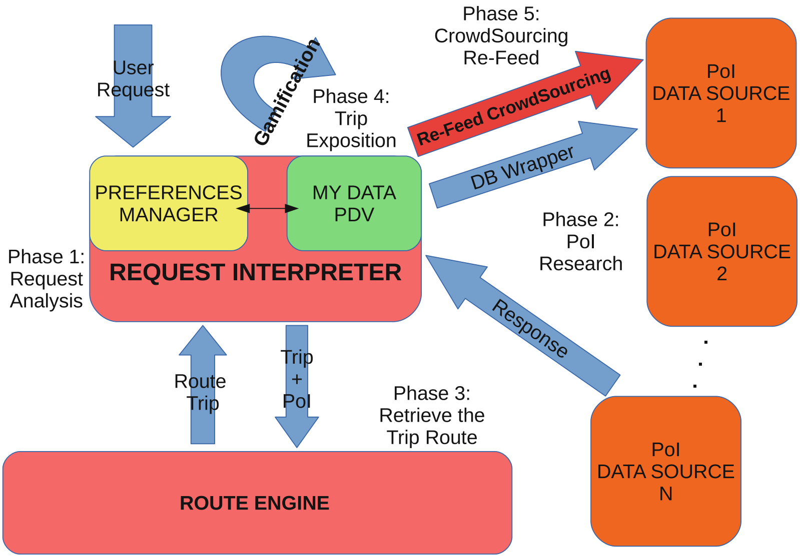


Fig. 1. Preference layer architecture

different and distinct phases [10]. The idea behind such a paradigm is that any function, that can be seen and managed as an independent service, have to be considered as a single microservice. A more complete service will be the result of a proper orchestration of those microservices.

The main advantage of this solution is that we can obtain an extremely modular and distributed architecture [9,11]. This kind of distributed compositionality becomes its greatest strength. As described in [12], even the internal components of the software are autonomous services, independent components conceptually deployed alone with dedicated technologies, both software and hardware. This means that this architectural style does not foster or forbid any particular programming, but it provides a partition structure of the component the developer has to follow. Since all the components of a microservice architecture are microservices, its distinguishing behavior derives from the composition and coordination of its components via messages. Thus, the core of a microservice platform will be the orchestration phase, namely the composition of microservices, tools, and processes invoked, and the connection and automation of workflows that create the final service [13]. However, the microservice paradigm introduces also new challenges and enlarges known threats. The distributed programming nature is at the same time its main advantage and disadvantage [14]. However it has been choose as development paradigm for the ease with which it was possible to disassemble the various stages in independent services. In this way, it could be potentially used later for other architectures. This preference layer interferes between the typical request of a common travel planner (e.g. OTP Analytics) and its response, intercepting the call and editing it based on a user preference processing. The architecture is structured around the orchestration of 5 main phases, which will be detailed in the following subsections:

1. Request and preference analysis.
2. Calculation of points of interest, and related gamification information on the various datasets.
3. Processing the route through a routing engine.
4. Presentation of the proposed route to the user and activation of the crowd-sourcing mode.
5. Post-processing of data received by the user and reintegration into the dataset.

2.1 Request Phase

Through applications (mobile and web-based ones) based on our architecture, the user can make a request for a generic path from a starting point S to a destination D. The user, as we will later show in the following section, has two possible choices. S/He can specify preferences for points of interest s/he wants to visit, or let the layer customization use preferences already stored in her/his profile.

In fact, our architecture follows the MyData Philosophy as regards the management of personal information of users [15]. MyData is a technology born in Finland whose purpose is to create an efficient and above all safe solution for the management and storage of personal data. The basic idea is to have a personal dataset called Personal Data Vault (hereafter PDV), where all the personal data of a user are properly stored. However, this dataset is accessed through a specific request, previously authorized, specifying needed data and format, so that the requesting service is not provided with rough and sensitive data, but only with the most appropriate result from a suitable data aggregation and processing algorithm. More details regarding this technology can be found in [16,17].

The Preferences layer (as shown in Fig. 1) then supports an authorized MyData engine, which, if no preference is specified, will be invoked to have the corresponding user profile. However, if the user specifies some preferences, (i) they are processed in a general way by the layer and then (ii) they are possibly integrated with the personal profile from the PDV. The purpose of these two phases is to generate and produce a single profile-request where the preferences are specified according to the format we use. This pre-processing step will then serve to calculate the various categories of preference and the weight to associate with each of them.

2.2 Points of Interest Search

Once the user's preferences are processed, the layer searches for the points of interest that match the request. It is necessary at this stage to better identify the data sources. Points of interest are calculated by invoking multiple data sources for several reasons. As stated earlier, the initial purpose of this project is to provide an architecture that can handle user's preferences, by driving him/her, within a generic urban area, among the monuments and other important points of interest in terms of cultural heritage.

We have specifically taken into account two main different types of data sources: official (coming from local entities and public administrations, such as municipalities, regions, ministries, departments, tourist offices, and so on) and unofficial ones (collected by means of crowdsourcing and crowdsensing activities, involving citizens and tourists).

In particular, as regards the official data sources, we have drawn our attention to the available resources related to the cultural heritage of the Emilia-Romagna Region (where Bologna is placed) as a case study. Thus, our system can communicate with the following official and authoritative data sources:

- the SACHER Project (http://www.eng.sacherproject.com/);
- Emilia-Romagna Cultural Heritage Open Data project (http://www.patrimonioculturale-er.it/webgis/);
- Bologna Open Map (http://dati.comune.bologna.it/bolognaopenmap?language=en);
- Google Maps.
- Open Street Map.

The DB wrapper is able to communicate with this Data Sources and provide a JSON structure, which is quite simple and includes the information needed to categorize monuments, palaces, museums or other points of interest related to cultural heritage, together with information about their geographical position and geo-referenced data. A fragment of the JSON code obtained from the search query, where geo-references are categorized by the type of monuments (according to the user's preferences) can be found in http://www.cs.unibo.it/~mirri/cicerone/json1.html.

Again, in this phase we have a twofold purpose. The first goal is obviously to look for points of interest, and this is accomplished by means of a deep search through all the datasets where the results are unified, eliminating duplicates and errors. Secondly, the list of corresponding matching POIs is analyzed, so as to identity those ones with missing metadata, related media and multimedia, additional relevant information.

This does not concern the geo-reference features, but documentation related to the specific point of interest and it includes data such as photos, timetables, accessibility for people with disabilities, real-time crowding, partial closures, etc.

The resulting POIs are indeed ordered and marked based on the missing features, so that they are recognizable once the path is calculated. At each missing feature is assigned a score that will be used on the fourth stage, so as to start the gamification phase. The final result is similar to the one shown at http://www.cs.unibo.it/~mirri/cicerone/json2.html. At this stage, the list of matching points of interest is ready to be exploited as an input to the route planning engine.

2.3 Routing Trip Elaboration

The bottom level planner depicted in Fig. 1 is an engine implementing routing algorithms that, given a geo-refined map and a possible additional data set (such as GTFS public transport data), computes a path from a starting point S to a destination D, going across or through all points of interest resulting from the previous step.

Any routing engine can be exploited for this step, in particular, we have used an OpenTripPlanner (OTP) instance in the prototype we have developed. The result of the OTP processing is another JSON file (a fragment is shown in http://www.cs.unibo.it/~mirri/cicerone/json3.html). It is worth noting that, at each point of interest, the necessary intermediate points are added to graphically map the path in the next step. Note that for each point, the corresponding ID is stored in such a way that it can memorize the association between the point and the information calculated in the previous steps, both in terms of interesting features of the POI, and of lack thereof (to involve the user in providing them).

2.4 User Trip

In the fourth step, the path is ready to be exploited by the user, following all the references made in the previous phases. At this stage, there is a first update of user's preferences within the corresponding PDV. Specifically, the computed path is saved and stored, together with all the relevant marks on the points of interest required and calculated. It is important to make a first update at this stage, rather than at the next ones because it is not certain whether the user will (want to) interact later.

The path is represented on a map (typically by using a mobile application or a web-based service, just like Google Maps or Open Street Maps); each POI with some missing features is marked and shown in the map so as to let the user be aware of the possibility of earning points by providing some of the missing information. Each feature then instantiates a listener, which basically his a web hook that starts and wait for possible interaction from the user. This interaction includes several possible different actions as:

- Add/Update a description or a previous one.
- Add/Update a picture or general media information or a previous one.
- Submit a review.
- Report a problem.

2.5 Crowdsourcing Re-feed

At this stage, the user can enjoy the required path with the relevant points of interest, add the missing data from the recommended points of interest, earn scores and gain bonus points. These data are again intercepted by the preference layer and will be used in two ways. Firstly, user's preferences and history will be updated a second time within the corresponding PDV. Secondly, the layer will feed the dataset by updating the new data derived from the user's interaction.

3 A Prototype

In order to prove our architectural approach and to conduct some preliminary tests, we developed a prototype web-app, which is designed to be exploited by web browsers and mobile devices. Such a web-app (named Cicerone) has been designed with the aim of supporting tourists, citizens, and visitors in wandering around the city of Bologna, across the points of interest related to cultural heritage, showing their details, and in expressing their preferences and personalizing their routes. It has been structured in back-end (developed by using NodeJS and MongoDB) and front-end (developed by using Vue.js, SASS, CSS3 Flexbox), as a responsive Web application. A screenshot of Cicerone user interface is depicted in Fig. 2.

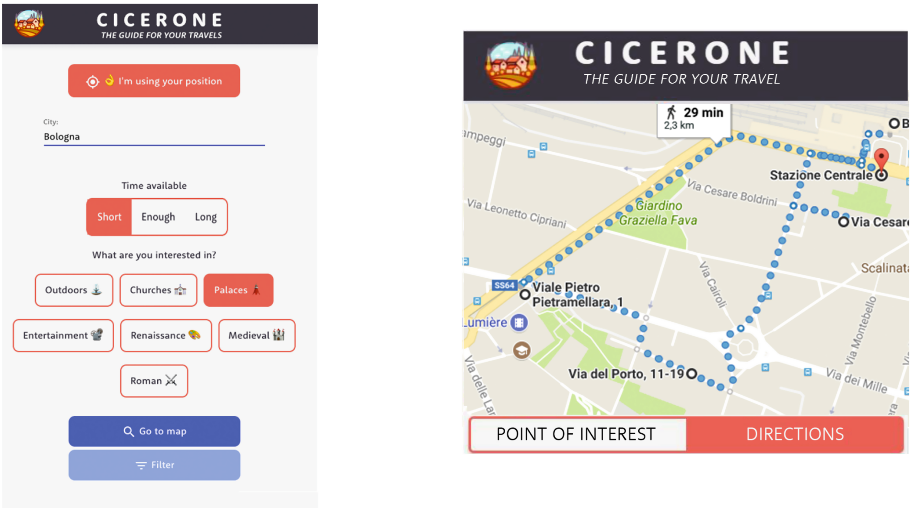


Fig. 2. A screenshot from Cicerone (left), John's path proposed by Cicerone (right)

On the basis of the described preference architecture, we have modified Cicerone app in order to be able to interact with such a preference layer. By means of Cicerone interface, the user can set some preferences in terms of time frame to spend during the visit, categories of favorite points of interests (i.e. type of monuments, ages, styles, etc.), starting position and destination of the path.

3.1 Personas and Scenarios

This subsection presents two personas and related scenarios we have defined with the aim of testing our prototype in the city of Bologna.

John stops over in Bologna during his travel by train from Venice to Florence. He has to wait 2 h for his connection, so he decides to use Cicerone to plan a small visit in Bologna.

As shown in Fig. 2, he uses the geolocation function, chooses “Short” in the time available radio button and selects that he’s interested in “Palaces” then clicks to “Go to map”. Cicerone proposes the following list of palaces to visit: (i) Palazzo delle Poste, (ii) Ex Conservatorio Margherita di Savoia, (iii) Ex Albergo Felsina, (iv) Palazzo Cantelli-Forti, (iv) Palazzo dell’Istituto di Aiuto Materno e di Assistenza ai Lattanti and (v) Ex Sede del Gruppo Rionale Fascista “Tabanelli”. The proposed visit, as shown in Fig. 2, takes about 30 min and starts, and ends, at Bologna train station. As many of the palaces are private and do not contain museums or exhibitions, Cicerone doesn’t consider a visit time for any point of interest, but it only proposes a personalized pedestrian path for such points [18]. During the path, Cicerone asks John to add some pictures of Palazzo delle Poste, since there is no photo available on the sources datasets.

Tomorrow morning, Jane will have a business meeting in Bologna, but she will have the afternoon off. She decides to have a walk to enjoy some of the churches in Bologna old town, using “Cicerone” to plan the visit. She enters Bologna in the city’s text field, choose “Enough” in the time available radio button and selects that she’s interested in “Churches” then clicks on “Go to map”. Cicerone proposes the following list of palaces to visit: (i) Chiesa di San Benedetto, (ii) Oratorio di San Carlo Borromeo, (iii) Ex Chiesa S. Maria del Buon Pastore, (iv) Chiesa dei SS. Filippo e Giacomo and (v) Ex Convento annesso alla Chiesa dei SS. Filippo e Giacomo. The visit proposed, as shown in Fig. 3, takes about 40 min and starts, and ends, at Bologna train station. At this time, however, the visit time for each point of interest has to be added. Actually the visit time is the same for each point of interest in “Church” category. During the visit Cicerone notifies some tags that can be added to the churches (e.g. accessibility for people with

Fig. 3. Jane’s path proposed by Cicerone

disabilities, availability of guided tours and visits, availability of wifi connection, availability of toilettes, religious services, services for babies, etc.). Cicerone lets the user also add information about the specific time devoted to the visit of each specific point of interest, enriching the whole system and improving the personalization settings and the computations related to the personalized routes.

4 Conclusions

In this paper, we presented a system, based on a microservice architecture, devoted to let the user set specific preferences and settings and to compute personalized routes across cultural heritage in a urban environment. The system integrates many and different data sources, mainly based on open data, provided by official entities (such as public and local administrations, e.g. municipalities, touristic offices, regions, etc.).

We are investigating crowdsourcing and gamification mechanisms that will enrich our system, with the aim of engaging users in providing and sharing additional data and resources related to the cultural heritage they meet during their routes. Thus, this turns citizens, visitors, and tourists into a kind of unofficial sources of information, which will be integrated with the official ones, improving the amount of available resources and supporting knowledge sharing activities. Another future work is based on the distribution of the proposed architecture in a cloud based one, by integrating in a deeper way the urban infrastructure [19], falling under the hat of Software as a Service (SaaS).

Acknowledgements. The authors want to thank Michele Campobasso, Alberto Giunta and Lisa Mazzini (who developed the first prototype of Cicerone), together with Francesca Galanti and Catia Prandi for their precious support and significant collaboration.

References

1. Marfia, G., Pau, G., Giordano, E., De Sena, E., Gerla, M.: VANET: on mobility scenarios and urban infrastructure. A case study. In: 2007 Mobile Networking for Vehicular Environments, pp. 31–36. IEEE (2007)
2. Mirri, S., Prandini, M., Salomoni, P., Delnevo, G., Callegati, F., Melis, A.: I want to ride my bicycle: a microservice-based use case for a MaaS architecture. In: 2017 IEEE Symposium on Computers and Communication (ISCC). IEEE (2017)
3. Bujari, A., Ciman, M., Gaggi, O., Palazzi, C.E.: Using gamification to discover cultural heritage locations from geo-tagged photos. Pers. Ubiquitous Comput. **21**(2), 235–252 (2017)
4. Dusi, N., Ferretti, I., Furini, M.: A transmedia storytelling system to transform recorded film memories into visual history. Entertain. Comput. **21**, 65–75 (2017)
5. Dusi, N., Ferretti, I., Furini, M.: PlayTheCityRE: a visual storytelling system that transforms recorded film memories into visual history. In: 2016 IEEE Symposium on Computers and Communication (ISCC), pp. 85–90, June 2016

6. Roccetti, M., Marfia, G., Palazzi, C.E.: Entertainment beyond divertissment: using computer games for city road accessibility. Comput. Entertain. (CIE) **9**(2), 10 (2011)
7. Oomen, J., Aroyo, L.: Crowdsourcing in the cultural heritage domain: opportunities and challenges. In: Proceedings of the 5th International Conference on Communities and Technologies, pp. 138–149. ACM (2011)
8. Stathopoulou, E., Georgopoulos, A., Panagiotopoulos, G., Kaliampakos, D.: Crowdsourcing lost cultural heritage. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **2**(5), 295 (2015)
9. Melis, A., Mirri, S., Prandi, C., Prandini, M., Salomoni, P., Callegati, F.: Integrating personalized and accessible itineraries in MaaS ecosystems through microservices. Mob. Netw. Appl. 1–10 (2017)
10. Mirri, S., Prandi, C., Salomoni, P., Callegati, F., Melis, A., Prandini, M.: A service-oriented approach to crowdsensing for accessible smart mobility scenarios. Mob. Inf. Syst. **2016** (2016)
11. Callegati, F., Gabbrielli, M., Giallorenzo, S., Melis, A., Prandini, M.: Smart mobility for all - a global federated market for mobility-as-a-service operators. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), pp. 1–8. IEEE (2017)
12. Newman, S.: Building Microservices. O'Reilly Media, Inc., Newton(2015)
13. Dragoni, N., Giallorenzo, S., Lluch-Lafuente, A., Mazzara, M., Montesi, F., Mustafin, R., Safina, L.: Microservices: yesterday, today, and tomorrow. CoRR, vol. abs/1606.04036 (2016)
14. Lea, G.: Microservices security: all the questions you should be asking (2015)
15. Kuikkaniemi, K., Poikola, A., Honko, H.: Mydata a nordic model for human-centered personal data management and processing. Technical report, Ministry of Transport Finland (2010)
16. Poikola, A., Kuikkaniemi, K., Honko, H.: Mydata - white paper. Technical report, November 2016
17. "Mydata 2016. Advancing human centric personal data," November 2016
18. Mirri, S., Prandi, C., Salomoni, P.: Personalizing pedestrian accessible way-finding with mpass. In: 2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC), pp. 1119–1124. IEEE (2016)
19. Callegati, F., Cerroni, W., Contoli, C., Santandrea, G.: Implementing dynamic chaining of virtual network functions in openstack platform. In: 2015 17th International Conference on Transparent Optical Networks (ICTON), pp. 1–4. IEEE (2015)

Privacy Preserving Multidimensional Profiling

Francesca Pratesi[1,2(✉)], Anna Monreale[2], Fosca Giannotti[1], and Dino Pedreschi[2]

[1] ISTI-CNR A. Faedo, Pisa, Italy
francesca.pratesi@isti.cnr.it
[2] Department of Computer Science, University of Pisa, Pisa, Italy
http://kdd.isti.cnr.it/homes/pratesi/

Abstract. Recently, big data had become central in the analysis of human behavior and the development of innovative services. In particular, a new class of services is emerging, taking advantage of different sources of data, in order to consider the multiple aspects of human beings. Unfortunately, these data can lead to re-identification problems and other privacy leaks, as diffusely reported in both scientific literature and media. The risk is even more pressing if multiple sources of data are linked together since a potential adversary could know information related to each dataset. For this reason, it is necessary to evaluate accurately and mitigate the individual privacy risk before releasing personal data. In this paper, we propose a methodology for the first task, i.e., assessing privacy risk, in a multidimensional scenario, defining some possible privacy attacks and simulating them using real-world datasets.

Keywords: Privacy risk assessment · Mobile phone data · Retail data

1 Introduction

One of the most pressing challenges of our time is understanding the complexity of our globally interconnected society: the big data originating from the *digital breadcrumbs* of human activities let us observe the ground truth of individuals at an unprecedented detail. Indeed, companies and governments are using this ocean of Big Data to unleash powerful analytic capabilities, connecting data from different sources, finding patterns and generating new insights. This can help transform the lives of individuals and help to solve many of society's challenges [23]. Hence, we want to explore a multidimensional scenario, providing an example of off-line analyses taking advantage of different data sources. The data can be purchasing data, movement tracks, electronic payments, social media data, and so on. The benefits of analyzing multidimensional data can be various, and they can lead to convenience for users, data providers or third parties.

Therefore, having the possibility to collect data from different sources, we want to analyze the implications from the point of view of both possible services

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 142–152, 2018.
https://doi.org/10.1007/978-3-319-76111-4_15

and privacy risk for participants. This means to address the study of quality and privacy in multidimensional data, i.e., data from different kind of sources. Clearly, each of these sources can lead to privacy leaks by itself, but if an attacker, who gain access to the datasets, has information about more aspects of her target's life (e.g., she knows some visited locations *and* some bought products), she can gain better chances to recognize her among all the users. The aim of this paper is to provide an example of a real-world privacy-preserving multidimensional service and to show what are the advantages and threats related to the use of such kind of data. Besides there are several strategies for the privacy risk mitigation, we illustrate the results using only the simplest solution, i.e., the suppression of risky individuals, showing that it is possible to obtain good results if we use the appropriate level of detail of the data.

The rest of the paper is organized as follows. In Sect. 2, we review some related work. In Sect. 3, we report the preliminaries of our work. In Sect. 4, we illustrate the service, the minimum data formats required for implementing it and the attacks we consider. In Sect. 5, we simulate the attacks and the results obtained at each step of the service. Section 6 concludes the paper.

2 Related Work

In recent years, privacy has been one of the most discussed issues. The aim of the methods proposed in the literature is assuring the privacy protection of individuals during both the analysis and the publishing of human data. The privacy has been studied in several contexts, from location based services [22] to GPS trajectories [1], from mobile phone data [2,20,24] to retail data [6,9,17]. In particular, our focus is on mobile phone and retail data. Privacy risks in mobile phone data, even in the case of releasing information with not fine granularity, are studied in [24]. In particular, authors consider the top N locations visited by each user. Typical solutions, highlighted by Blondel et al. [2] are to operate small modification of datasets or to change frequently (at least daily) pseudo-identifiers. Unfortunately, this can lead big limitation on analyses and services that can be performed. In [8], it is suggested to use synthetic data, which can reproduce many features of mobility of users. An extension of this work that relies on Differential Privacy can be found in [11]. Retail data are analyzed for maximizing the profit of companies [13] and for potentiating the customer care [18], but they might reveal personal or sensible information about the subject (for example, if he discovers that the subject regularly buys gluten-free pasta, it is easy to infer that the user suffers from celiac disease). To prevent these issues, researchers have developed privacy preserving methodologies, in particular, to extract association rules from retail data [6,9,17].

In the last year, different techniques for risk management[1] has been proposed, such as the OWASP's [12] and LINDDUN [4] methodologies, Microsoft's DREAD [10], SEI's OCTAVE [7]. Unfortunately, many of them do not consider in deep

[1] The risk evaluation task is compliant with the EU General Data Protection Regulation.

privacy. A methodology that presents an approach for the systematic privacy evaluation is the one presented in [16] and extended in [14,15]. This represents a starting point for our work and it will be outlined in Sect. 3.

3 Privacy Risk Assessment Framework

In this paper, we consider the work proposed in [16], which allows for the privacy risk assessment of human data, considering a scenario where a Service Developer asks a Data Provider for data to develop an analytical service. The Data Provider must guarantee the right to privacy of the individuals whose data are recorded. In a nutshell, the framework PRISQUIT is based on the Privacy by Design paradigm, so the first specification is to establish the data requirements for the service. Then the Data Provider queries its dataset, producing a set of datasets with different data structures and aggregations and it then: *(i)* identifies the background knowledge that an adversary might have about her target; *(ii)* simulates the attack based on that background knowledge, computing the privacy risk values for every individual; *(iii)* selects the dataset with the best privacy-utility trade-off; *(iv)* applies a privacy risk mitigation method (e.g., generalization, randomization, suppression) on that dataset; and *(v)* delivers the sanitized dataset to a third party.

In [16], and consequently in this paper, it has been used as privacy risk the risk re-identification [19], whose related attacks assume that an adversary gains access to a dataset and, using some background knowledge about an individual under attack, he/she tries to re-identify that individual in the dataset. The background knowledge represents both the kind and quantity of information known by the adversary. We use b to indicate the specific background knowledge (e.g., the fact that a user visited a certain location on a certain day) and B_h to indicate a set of background knowledge of size h (e.g., B_2 can represent all the possible couple of locations visited by an individual).

Let $\mathcal{D}$ be a database, D a dataset derived from $\mathcal{D}$ (e.g., an aggregated data structure on time and/or space), and D_u the set of records representing a user u in D, the probability of re-identification is defined as follow.

Definition 1 (Probability of re-identification *[16]***).** *Given an attack, a function matching(d, b) indicating whether or not a record $d \in D$ matches the background knowledge b, and a function $M(D, b) = \{d \in D | matching(d, b) = True\}$, we define the probability of re-identification of an individual u in dataset D as: $PR_D(d = u|b) = \frac{1}{|M(D,b)|}$ that is the probability to associate record $d \in D$ to individual u, given background knowledge b.*

Note that $PR_D(d{=}u|b) = 0$ if the user u is not in D. Since each background knowledge b has its own probability of re-identification, we define the risk of re-identification of an individual as the maximum probability of re-identification over the set of possible background knowledge:

Definition 2 (Privacy risk *[16]*). The risk of re-identification (or privacy risk) of an individual u given a set of background knowledge B_k is her maximum probability of re-identification $Risk(u, D) = \max PR_D(d = u|b)$ for $b \in B_k$. It has the lower bound $\frac{|D_u|}{|D|}$ (a random choice in D), and $Risk(u, D) = 0$ if $u \notin D$.

An individual is hence associated with several privacy risks, each for every background knowledge of an attack.

4 Privacy Preserving Multidimensional Profiling

In [16], we have a scenario where a single Data Provider (DP) and a single Service Developer (SD) interact, using PRISQUIT, to determine the best dataset to be released, regarding the trade-off between privacy and utility. In this paper, instead, we have an ecosystem where several DPs interact with an SD, through a Safe and Trusted Environment (STE) for personal data collection and sharing, which provides a link between DPs and SD and assisting DPs in the use of PRISQUIT. Indeed, the SD asks to the STE for a series of data, and they define together the minimum data format suitable to provide a reliable service. The STE address the study of quality and privacy in multidimensional data, i.e., data from different kind of sources. In order to have a clear view of both the actual privacy risk and the potentiality of services with combined datasets, we should consider that users appear in each source. Only by relying on this basis we can provide a rigorously correct quantification. However, we can analyze a what-if scenario, where users belonging to different datasets are linked to each other, pretending to be the same individual. In the following, we outline a possible service that needs multidimensional sources, analyzing the minimum data format necessary and the attacks that is possible to perform on every single dataset separately or exploiting the presence of same users in more datasets.

4.1 Promotion Service Based on Recurrent Events

We suppose that a marketing manager of a retail chain wants to suggest products related to a recurrent event, such as an annual movie festival or a monthly organic market, to her customers. The marketing expert decides to target the campaign only to customers that actually participated in the event in the past since they will be more likely to respond positively to the advertisements.

In order to implement this service, we need to have access to a dataset from a telco operator to obtain the participants and one from a retail chain to verify their purchases. Indeed, the service is composed of the following steps.

Event Detection and Users Participation. The SD asks to the STE for mobile phone data. Then, the SD can detect relevant peaks representing an event, comparing the density of population within a region in a given moment against the expected density for that area [21]. The SD can also discover the users that participated in that event, discriminating between actual visitors and regular inhabitants [5].

Marketing Campaign Definition. Then, the SD can ask for the purchases, related to a specific time window before the event, of the participants. The size of the time window is strictly related to the event: it could be a day or a week time window, or even a monthly one if the event needs some preparation, like Halloween or Christmas. The SD defines what the products related to the event are, and he checks if users bought those products. If not, the SD can contact the retail chain marketing responsible (through the STE, which is responsible for the match of users among the different data sources), suggesting how to spread a targeted campaign before the following occurrence of the event.

Minimum Data Formats

Since we rely on the data minimization principle[2] and on the Privacy-by-Design model [3], we define the minimum data format necessary to develop this service.

Minimum Data Format (Mobile Phone Data). We want to obtain a density map of a territory, so we only need an aggregation like the one presented in [5]. Therefore, the data format is a profile, i.e., a matrix P where i denotes the time slot (e.g., morning, afternoon, night) and j denotes the day (if the event is a daily one, we will remove the time slots). Since we do not need the precise activities of users, the profile contains only 1 (indicating a presence) and 0 (absence).

$$P^u_{ij} = \begin{cases} 0 & \text{if } u \text{ is not present or he performs no calls at day } j \text{ and slot } i \\ 1 & \text{if } u \text{ performs at least one call in the area at day } j \text{ and slot } i \end{cases} \tag{1}$$

Minimum Data Format (Purchasing Data). Here, the data format is the set of distinct items bought in a determined time window:

$$\{i_1, i_2, \dots i_n\} \tag{2}$$

Clearly, considering a product at the detail level of the bar code is usually unnecessary, so we can climb the product taxonomy and establish the essential level of detail, like categories or subcategories, suitable for the service realization.

Attack Models

We point out that each attack assumes the adversary gains access to the dataset. Performing an attack means finding a set C of possible matches for a target, given a certain background knowledge. The probability of re-identification of the user u is $\frac{1}{|C|}$. A greater number of candidates implies a better privacy protection.

Background Knowledge on Mobile Phone Data. We suppose that the attacker knows part of the data format (1), e.g., she knows that her target was in Florence on Monday and Friday of a specific week. Thus, the adversary can build a partial (but exact) profile b, where $b_{ij} = -1$ if the attacker does not have any

[2] Art. 5 EU GDPR.

information about the period (i, j). We suppose that h represents the number of weeks in which the attacker knows information about the calls of her target, so B_h corresponds to the possible combination of weeks of the profile P.

Attack on Mobile Phone Data. The attacker uses the background knowledge b on the user u to match all the profiles that include b. The set of matched profiles is $C = \{P \in \mathcal{P} | \forall b_{ij} \geq 0. b_{ij} = P_{ij}\}$.

Background Knowledge on Retail Data. We suppose that the attacker has as background knowledge a subset of products bought by her target, in the format (2); for example, the attacker once saw the shopping chart of her target. The subset of items known has size q, so we have as background knowledge $b \in B_q$.

Attack on Retail Data. The attacker uses the background knowledge b on the user u to match all the set of items that include b. Given $D(u_i)$ the set of items of the user $u_i \in \mathcal{D}$, the candidate set is computed as $C = \{u_i | b \subseteq D(u_i)\}$.

In the following, we define a possible multidimensional background knowledge and the correspondent attack. Note that, in order to execute this attack, it is necessary to have a link between users: for example, we could know that the caller 100 in the mobile phone dataset is the customer 30 in the retail dataset.

Background Knowledge on Mobile Phone and Retail Data. Here, we suppose that the attacker has as background knowledge a subset of call activities and products bought by her target, i.e., a combination of the two previous attacks. The subsets of items and call activities may have variable size q and h, respectively.

Attack on Mobile Phone and Retail Data. Given the sizes q and h, we denote by B_q and B_h the two set of background knowledge. Let $D_p(u_i)$ the set of items of the user u_i and $D_m(u_i)$ the weeks of call activities of u_i in the format (2) and (1), respectively. For each instance b' of B_q and b'' of B_h, the candidate set is computed as: $C = \{u_i | b' \subseteq D_p(u_i) \wedge b'' \subseteq D_m(u_i)\}$ The probability of re-identification given the background knowledge $b = \{b' \cup b''\}$ is $\frac{1}{C}$.

5 Experiments

Here, we present the simulation of the attacks presented in Sect. 4. Firstly, we illustrate the datasets; then, we provide the results of the simulation of privacy risk and an indication of the output service-side, i.e., from the SD's point of view. We provide the simulations for all the attacks; however, since we do not have the actual link between users in the two datasets, we report a what-if analysis for the multidimensional attack, imagining that the link between callers and customers is known by the STE. With this aim, we randomly selected customers, and we forcibly assigned them to some of the callers (in particular, we assumed that STE knows the purchases of one-third of the mobile phone population).

5.1 Datasets Presentation

The mobile phone dataset is provided by one of the major Italian mobile operators. It regards the territory of a great part of Tuscany (106 municipalities out of 276), for a period from February 17, 2014, to March 23, 2014, and it reports the activities of around 858k individuals, for a total of 51 million call records. The retail dataset is provided by Unicoop Tirreno, one of the major retail distribution companies in Italy. It regards a time window spanning from January 1, 2007, to June 30, 2014. The active and recognizable customers, i.e., users with a loyalty card and with at least a purchase in the period, are about 800,000.

5.2 Privacy Analysis

As illustrated in Sect. 4, we start from mobile phone data. The SD asks for this kind of data among the Tuscan municipalities, in order to detect events. The STE builds the profiles aggregated in days (i.e., from midnight to midnight) of all the people present in each area. At a certain point, the STE analyzes the Viareggio municipality, i.e., it simulates the attack on mobile phone data presented in Sect. 4.1, obtaining the cumulative distribution of probabilities of the privacy risk shown in Fig. 1(a). Here, we can find the percentages of users having at most a certain risk of re-identification, obtained varying the background knowledge from 1 to 4 weeks. As we can see, due to the aggregated nature of the data, if we hypothesize that the attacker knows 1 week of calls of her target, the probability she succeeds in re-identification is extremely low, since for 95% of users the privacy risk is below 0.005 (i.e., they are indistinguishable from at least 199 others). As we can expect, if we increase the number of weeks known by the attacker, the risk increases too. However, it is never dramatically high: knowing 2 weeks, we have that only 10% of users have a risk greater than 0.2, while knowing 3 weeks this risk is associated with 35% of individuals. It is interesting to note the "knee" in the curves: its presence indicates that until that point it is possible to obtain a lower risk renouncing to relatively few users.

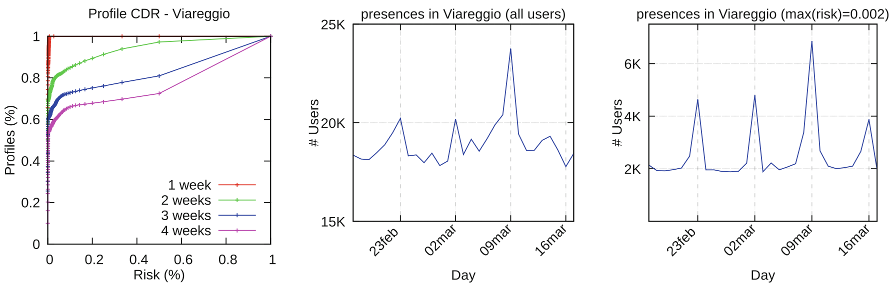


Fig. 1. (a) CDF of the privacy risk, varying the background knowledge. (b–c) The timeseries of presences in Viareggio, using mobile phone data: (b) timeseries computed with all the original users; (c) timeseries computed with only safe users, fixing the maximum risk of re-identification to 0.002

Since we are mainly interested in visitors, i.e., exceptional presences w.r.t. the routine, we can imagine that they are individuals present only for the event.

So, they are less problematic concerning privacy w.r.t. commuters and residents, since, likely, they are naturally similar to other visitors. For this reason, we can decide to use a quite extreme privacy threshold, expecting more or less the same efficacy in terms of quality. In particular, we chose to guarantee a maximum risk of re-identification among users of 0.002 (that corresponds to a group size of at least 500), releasing 48% of profiles (i.e., 41K users out 86K).

As soon as the SD receives these profiles, it performs a peak detection, analyzing the presences, day by day, in the considered month. In order to provide a fair comparison of the utility, we show in Fig. 1(b) the results of the peak detection analyzing the profiles of all the 86 K users (so, *without any regards for privacy*) and in Fig. 1(c) the same analysis applied only to released data (i.e., the profiles of users with the maximum risk of re-identification equal to 0.002). Even if the scale between the two plots differs from almost an order of magnitude, we can clearly recognize peaks on Sundays. In particular, Fig. 1(b) and (c) have an anomalous peak on Sunday, March 9, so the SD presumes that something different occurred that day. Indeed, on that Sunday there was a popular Carnival Parade. Using data depicted in Fig. 1(c), the SD can also discriminate between people that regularly live the area (i.e., residents and commuters), who represent the baseline of about 2,000 persons, and real visitors, who define the peak. The SD is interested in asking for the expenses of these 3,500 users (the peak is around 5,000 users higher than the average, but we can assume, for the sake of simplicity, that 1,500 individuals visited Viareggio every Sundays, so we remove them from the count). At this point, we rely on the hypothesis that the STE can have access to the purchases of one-third of these individuals, so we randomly select 1,000 customers from the retail dataset.

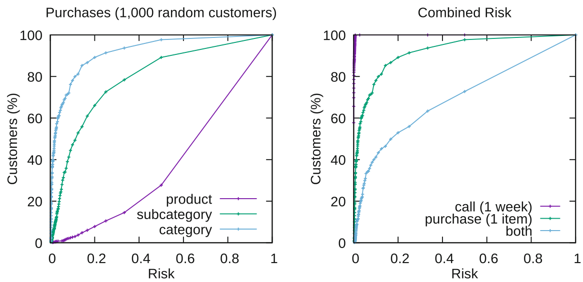


Fig. 2. Simulation of the attack on retail data varying the level of detail of items (Left) and of the attack on call activity and retail data using a random matching between callers and customers (Right)

The second step of the analysis is assessing the privacy risk of these 1,000 customers. STE analyzes the purchases in a time window covering the period in which the supermarket chain sells products related to Carnival, i.e., from January 1 to March 9, 2014. The risk of re-identification of these users, fixing the number of known products $h = 1$, is reported in Fig. 2 (Left). From this plot, it is immediate to see that using a detailed level leads to a too high risk. However, if we use less detailed information, we obtain better results. In order to clarify the taxonomy, we provide a simple example describing the three levels of detail: we have a specific milk (e.g., "500 ml UHT skim milk of brand Coop") at the *product* level; as *subcategory* we have "UHT" or "fresh" milk; at the *category* level we only know that the item is "milk". In our case, at the category level, we can still search for Carnival related items. Thus, releasing the categories, we have that 90% of users have 0.25 as maximum risk, while 78% have 0.1 and 65% have 0.05. The STE can decide to release the purchases of the customers having a maximum risk of 0.05, i.e., who are in groups of at least 20 indistinguishable users. The released customers are 655. The SD analyzes their purchases and discovers that only 37 of them bought in Unicoop stores products related to the event. This means that SD can inform Unicoop that, among its customers, 618 individuals very likely participated in Viareggio Carnival in 2014, but they did not buy anything in the month preceding the event; maybe with focused offers, they will be inclined to buy in Unicoop stores before the following Carnival event.

Lastly, if we consider the multidimensional background knowledge attack on mobile phone data and retail data, which is based on a very strong knowledge, we obtain the simulation depicted in Fig. 2 (Right). Here, we can see that, even if knowing only one aspect of the users' life does not compromise the possibility to release a portion of data without high privacy concerns, when we combine the two dimensions and the certainty to know who is each individual in the two datasets, the risk of re-identification is substantially higher. Indeed, if we want to provide the same maximum risk level we used before, i.e., 0.05, we can release only 299 individuals' data.

6 Conclusion

In this paper, we envision a data sharing scenario, where a trusted component called Safe and Trust Environment (STE) offers an interface between different Data Providers (DP) and potential Service Developers (SD). The STE receives queries by SDs and asks for the correspondent data to specific DPs, defining the minimum data format necessary and analyzing the privacy risk of each individual whose data belong, *before releasing the data*. We defined some attacks related to two kinds of data, in order to analyze the risk of sharing data with some background knowledge about them both separately and together. We simulated these attacks on real-world datasets, discovering that using the minimum information required it is possible to achieve a good trade-off between privacy protection and quality of service. However, combining little information from different sources can lead to risk sensibly higher than a larger amount of knowledge related only to one source. As future work, we would analyze different services, including other kinds of data, like GPS tracks or credit card logs, in order to provide realistic examples of privacy-preserving multidimensional data sharing.

Acknowledgment. Funded by the European project SoBigData (Grant Agreement 654024).

References

1. Abul, O., Bonchi, F., Nanni, M.: Anonymization of moving objects databases by clustering and perturbation. Inf. Syst. **35**(8) (2010)
2. Blondel, V.D., Decuyper, A., Krings, G.: A survey of results on mobile phone datasets analysis. EPJ Data Sci. **4**(1), 10 (2015)
3. Cavoukian, A.: Privacy by design the 7 foundational principles, August 2009
4. Deng, M., Wuyts, K., Scandariato, R., Preneel, B., Joosen, W.: A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requir. Eng. **16**(1) (2011)
5. Gabrielli, L., Furletti, B., Trasarti, R., Giannotti, F., Pedreschi, D.: City users' classification with mobile phone data. In: IEEE Big Data 2015 (2015)
6. Giannotti, F., Lakshmanan, L.V., Monreale, A., Pedreschi, D., Wang, H.: Privacy-preserving mining of association rules from outsourced transaction databases. IEEE Syst. J. **7** (2013)
7. Institute, C.S.E.: Octave. http://www.cert.org/octave/
8. Isaacman, S., Becker, R., Cáceres, R., Martonosi, M., Rowland, J., Varshavsky, A., Willinger, W.: Human mobility modeling at metropolitan scales. In: MobiSys 2012 (2012)
9. Le, H.Q., Arch-int, S., Nguyen, H.X., Arch-int, N.: Association rule hiding in risk management for retail supply chain collaboration. Comput. Indus. **64** (2013)
10. Meier, J., Corporation, M.: Improving Web Application Security: Threats and Countermeasures. In: Patterns & Practices, Microsoft (2003)
11. Mir, D.J., Isaacman, S., Cáceres, R., Martonosi, M., Wright, R.N.: Dp-where: differentially private modeling of human mobility. In: IEEE Big Data 2013 (2013)
12. OWASP: Risk rating methodology. https://www.owasp.org/index.php/OWASP_Risk_Rating_Methodology
13. Pauler, G., Dick, A.: Maximizing profit of a food retailing chain by targeting and promoting valuable customers using loyalty card and scanner data. EJOR **174** (2006)
14. Pellungrini, R., Pappalardo, L., Pratesi, F., Monreale, A.: A data mining approach to assess privacy risk in human mobility data, ready to appear in ACM TIST
15. Pellungrini, R., Pratesi, F., Pappalardo, L.: Assessing privacy risk in retail data. In: PAP@ECML-PKDD 2017 (2017)
16. Pratesi, F., Monreale, A., Trasarti, R., Giannotti, F., Pedreschi, D., Yanagihara, T.: Prisquit: a system for assessing privacy risk versus quality in data sharing, Technical report 2016-TR-043. ISTI - CNR, Pisa, Italy (2016)
17. Rizvi, S.J., Haritsa, J.R.: Maintaining data privacy in association rule mining. In: VLDB 2002 (2002)
18. Rygielski, C., Wang, J.C., Yen, D.C.: Data mining techniques for customer relationship management. Technol. Soc. **24** (2002)
19. Samarati, P., Sweeney, L.: Generalizing data to provide anonymity when disclosing information (abstract). In: PODS, p. 188 (1998)
20. Song, Y., Dahlmeier, D., Bressan, S.: Not so unique in the crowd: a simple and effective algorithm for anonymizing location data. In: PIR@SIGIR 2014 (2014)
21. Trasarti, R., Olteanu-Raimond, A.M., Nanni, M., Couronné, T., Furletti, B., Giannotti, F., Smoreda, Z., Ziemlicki, C.: Discovering urban and country dynamics from mobile phone data with spatial correlation patterns. Telecommun. Policy **39**(3–4) (2015)

22. Xiao, Y., Xiong, L.: Protecting locations with differential privacy under temporal correlations. In: ACM CCS 2015 (2015)
23. World Economic Forum: Rethinking personal data: Strengthening trust. http://www3.weforum.org/docs/WEF_IT_RethinkingPersonalData_Report_2012.pdf
24. Zang, H., Bolot, J.: Anonymization of location data does not work: a large-scale measurement study. In: MobiCom. ACM (2011)

Fall Detection with Kinect in Top View: Preliminary Features Analysis and Characterization

Susanna Spinsante(✉), Manola Ricciuti, Enea Cippitelli, and Ennio Gambi

Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, 60131 Ancona, Italy
s.spinsante@staff.univpm.it
http://www.tlc.dii.univpm.it

Abstract. Fall detection is a well investigated research area, for which different solutions have been designed, based on wearable or ambient sensors. Depth sensors, like Kinect, located in front view with respect to the monitored subject, are able to provide the human skeleton through the automatic identification of body joints, and are typically used for their unobtrusiveness and inherent privacy-preserving capability. This paper aims to analyze depth signals captured from a Kinect used in top view, to extract useful features for the automatic identification of falls, despite the unavailability of joints and skeleton data. This study, based on a set of signals captured over a number of test users performing different types of falls and activities, shows that the speed of falling computed over the blob identifying the person, extracted from the depth images, should be used as a feature to spot fall events in conjunction with other metrics, for a better reliability.

Keywords: Fall detection · Depth image processing · Blob Features · Speed of falling

1 Introduction and Background

In recent years, many consumer electronics devices and products, like traditional appliances, have found new and sometimes unexpected adoption as different tools, usually as sensors. In fact, in the broadest definition, a sensor is "*an electronic component, module, or subsystem whose purpose is to detect events or changes in its environment and send the information to other electronics, frequently a computer processor*" [7]. With respect to this definition, a common appliance, like a fridge, once turned into a *smart* object, may become a sensor to detect and analyse events related to Activities of Daily Living (ADLs) like cooking or having meals, and enable an unobtrusive behavioral analysis to recognize anomalous habits [4,5]. The same happened with a commercial product designed and shipped for gaming purposes by Microsoft, the Kinect, a motion

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 153–162, 2018.
https://doi.org/10.1007/978-3-319-76111-4_16

sensing input device for video game consoles and Microsoft Windows PCs. Based around the concept of a webcam-style add-on peripheral, Kinect enables users to control and interact with their console or computer without the need for a game controller, just using gestures and spoken commands, i.e. through a natural user interface. Since Kinect became fully available for PC users, together with its Software Development Kit (SDK), many researchers started using it for applications not related to gaming, but pertaining to gesture, action and activity recognition and computer vision, in a broad sense, spanning from device-free interaction with other systems or devices, to gait and posture analysis, to fall detection and remote rehabilitation. Examples of the aforementioned applications may be extensively found in the literature (see [1,3,6], among others).

Among the possible fields of application of the Kinect device, this paper investigates its use in fall detection, assuming a specific constraint. In fact, differently from most of the literature available on this topic, which assumes to use the sensor in a front view configuration, we use Kinect in a top-view setting, i.e. the sensor is installed on the ceiling of a room (like the lab into which tests and experiments have been carried out). This way, the subject's skeleton, and the joints' coordinates Kinect is able to compute when used in front view, are not available for processing. In a previous paper [2], we presented an algorithm for automatic fall detection exploiting the Kinect in the same top-view configuration. That work based the fall detection capability on a number of depth image processing functions, able to recognize the blob of a person and check its geometric features against a number of anthropometric thresholds, and on comparing the relative height of the blob with respect to the floor (details are available in [2]).

In this paper, we extensively test the algorithm on different types of falls, to identify a suitable feature (generated by processing the blob data) for the purpose of automatic fall detection and classification. In fact, we focus on the speed of movement of the person's blob during the fall, and on the amount of variation of the blob's height with respect to the floor, to check if these figures may be used to discriminate between a fall or an ADL, or among different types of falls. Again, several works in the literature build upon this idea, but they exploit the subject's head joint coordinates provided by the sensor used in a front-view configuration, or the center of mass coordinates computed from the joints. We opt for a top-view setup, as it results to be far less obtrusive in a real-life scenario, and even more robust to possible occlusions due to objects, like furniture, located in the monitored environment.

The paper is organized as follows: Sect. 2 presents both materials and methods used in our study, whereas Sect. 3 describes and discusses the experiments carried out and the results obtained. Finally, Sect. 4 concludes the paper.

2 Materials and Methods

The system setup adopts a Kinect device in top view configuration, at a maximum height of 3 m from the floor, thus providing a coverage area of 8.25 m^2. To extend the monitored area, in principle the sensor could be elevated up to around 7 m (far higher than typical living environments); beyond this distance the depth data become unreliable [8]. The algorithm described in [2] works with raw depth data (given in millimeters), that are captured at a frame rate of 30 fps with a resolution of 320×240 pixels, using the Microsoft SDK v.1.5. The depth signal output by the sensor is filtered to reduce noise and to improve the blob identification. Once the person's blob is identified, the algorithm assigns it a *centroid*, the 3D coordinates of which are stored during the activity execution. The fall detection tool identifies the falls and uses different colors (yellow, red, green) to notify conditions of warning, fall and recovery, respectively. In this work, we use the 3D coordinates of the centroid and process them to obtain its speed of movement along the vertical direction (z axis).

2.1 The Fall Detection Algorithm

The fall detection algorithm, the details of which have been published in [2], relies upon some basic operations that are schematically shown in Fig. 1 and can be summarized as follows:

- preprocessing and segmentation: the incoming depth frame is pre-processed to enable the succeeding steps, and a so-called *reference frame* is generated to improve the identification of human subjects;
- distinguish object step: this component of the algorithm identifies, splits, and classifies into objects the different clusters of pixels in the reference frame;
- identification of human subject: starting from the set of separated objects, those representing a human subject are identified by checking several anthropometric relations;
- subject tracking and fall detection: the system tracks the movements of the human subjects in the depth frames, and detects if a fall occurs. The possible fusion of blobs occurring when two or more subjects get in contact is properly handled.

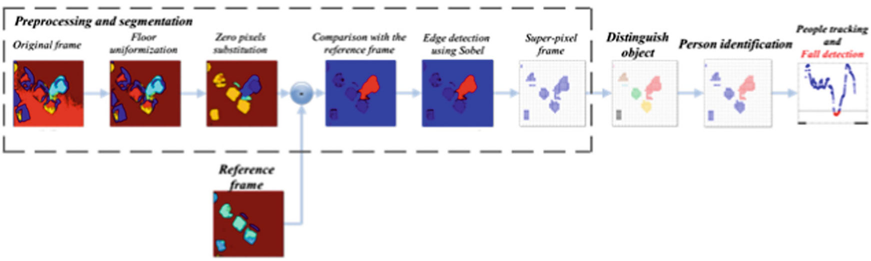


Fig. 1. Overview of the fall detection system (from [2])

The algorithm, conceived and designed with low complexity in mind, identifies the falls by comparing the depth values taken by the pixels representing the human subject's centroid to a threshold empirically set to 400 mm from the floor. When the detected height of the centroid undergoes the threshold, the algorithm triggers an alarm to notify a fall event. As shown in Fig. 2, the fall detection implemented by the algorithm depends only on the relative height of the centroid to the threshold, not on its time variation along the vertical axis. By exploiting the data provided by the algorithm (the coordinates of the centroid), and based on the knowledge of the frame rate, in this work we investigate the relation between the time variation of the centroid vertical position (i.e. its velocity along the z axis) and the type of action (fall or ADL) simulated by the subject, to possibly identify a useful feature for automatic action/fall recognition independent from the threshold value that needs to be heuristically defined.

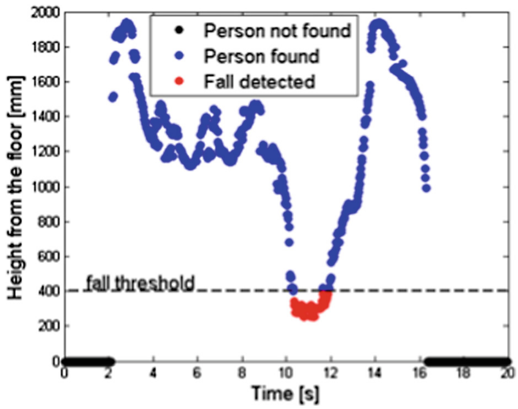


Fig. 2. Sample sequence of depth values taken by the subject's centroid, compared to the 400 mm threshold. A fall event is detected and highlighted in red (Color figure online)

2.2 Experimental Protocol

In order to evaluate the algorithm performances, several tests were carried out in a laboratory environment. Tests involved 17 healthy subjects, 3 females and 14 males, aged between 21 and 55 years. More specifically, 32 different types of falls and 8 ADLs have been simulated by each tester, for a total amount of 544 falls and 136 ADLs. The experimental protocol is reported in Table 1. The simulated falls can be divided into 8 groups: backward finishing lying, backward finishing sitting, forward finishing lying, forward finishing on knees, forward finishing on knees grabbing a chair, forward finishing on knees grabbing a sofa, left side, and right side. For each of them, 4 different situations have been considered: (i) falling from the stand position and then remaining on the ground; (ii) falling from the stand position and then recovering; (iii) falling during walking and then remaining on the ground; (iv) falling during walking and then recovering.

Regarding ADLs, the performed activities can be divided into 4 groups: pick up an object from the floor with bending, pick up an object from the floor with squatting, sit and get up from a chair, and sit and get up from a couch. For each of them, two situations have been considered: (i) from a standing position; (ii) from walking.

All testers were asked to sign an informed consensus form before starting the experiment, and let perform the tasks freely. Foam mats were used to soften the blow and protect subjects from injuries when simulating falls.

3 Results and Discussion

As a first analysis, we considered falls belonging to the classes FBESFR (backward fall, finishing sitting), FFOKFR (forward fall on the knees), FFOKSO (forward fall, on the knees, grabbing the sofa), and FFOKCH (forward fall, on the knees, grabbing the chair). Sample trajectories of the four classes in the 3-D space are shown in Figs. 3(a)–(d), for one of the subject who performed the experimental tests. As visible in the Fig. 3(a) and (b), a fall belonging to class FBESFR originates a different trajectory than a fall belonging to class FFOKFR, notably featuring an opposite direction. On the other hand, falls belonging to classes FFOKSO and FFOKCH (Fig. 3(c) and (d)) have similar patterns and differ for the value of the z coordinate in the final position (denoted by t_end).

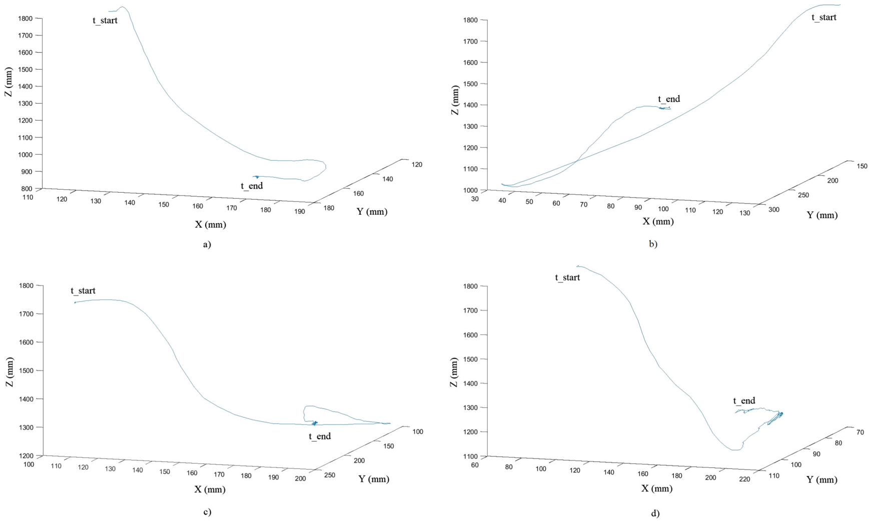


Fig. 3. Sample fall trajectories for the classes (a) FBESFR, (b) FFOKFR, (c) FFOKSO, and (d) FFOKCH for one of the subjects executing the test. Labels t_start and t_end denote the starting and ending point of the trajectory

For each type of fall within each class, the speed of fall along the z axis has been computed as the difference between the z coordinate value in two consecutive frames, captured at a rate of 30 fps, i.e. over a 33 ms time interval.

Table 1. Summary of fall detection experiments. "Subj." stands for subject.

Type	Activity name	Description
Backward fall, finishing lying	FBELFRST	Subj. is standing, falls backwards, and remains on the ground
	FBELFRSTRC	Subj. is standing, falls backwards, stays on the ground for a while and then gets up again
	FBELFRWK	Subj. walks, falls backward, and remains on the ground
	FBELFRWKRC	Subj. walks, falls backward, stays on the ground for a while and then gets up again
Backward fall, finishing sitting	FBESFRST	Subj. is standing, falls backwards, and remains on the ground
	FBESFRSTRC	Subj. is standing, falls backwards, stays on the ground for a while and then gets up again
	FBESFRWK	Subj. walks, falls backward, and remains on the ground
	FBESFRWKRC	Subj. walks, falls backward, stays on the ground for a while and then gets up again
Forward fall, finishing lying	FFELFRST	Subj. is standing, falls forwards, and remains on the ground
	FFELFRSTRC	Subj. is standing, falls forwards, stays on the ground for a while and then gets up again
	FFELFRWK	Subj. walks, falls forwards, and remains on the ground
	FFELFRWKRC	Subj. walks, falls forwards, stays on the ground for a while and then gets up again
Forward fall on the knees grabbing the chair	FFOKCHST	Subj. is standing, falls forwards, and remains on the ground, grabbing the chair
	FFOKCHSTRC	Subj. is standing, falls forwards, stays on the ground grabbing the chair for a while, and then gets up again
	FFOKCHWK	Subj. walks, falls backward, and remains on the ground, grabbing the chair
	FFOKCHWKRC	Subj. walks, falls backward, stays on the ground grabbing the chair for a while, and then gets up again
Forward fall on the knees	FFOKFRST	Subj. is standing, falls forwards, and remains on the ground
	FFOKFRSTRC	Subj. is standing, falls forwards, stays on the ground for a while and then gets up again
	FFOKFRWK	Subj. walks, falls forwards, and remains on the ground
	FFOKFRWKRC	Subj. walks, falls forwards, stays on the ground for a while and then gets up again
Forward fall on the knees grabbing the sofa	FFOKSOST	Subj. is standing, falls forwards, and remains on the ground, grabbing the sofa
	FFOKSOSTRC	Subj. is standing, falls forwards, stays on the ground grabbing the sofa for a while and then gets up again
	FFOKSOWK	Subj. walks, falls forwards, and remains on the ground, grabbing the sofa
	FFOKSOWKRC	Subj. walks, falls forwards, stays on the ground grabbing the sofa for a while and then gets up again
Left side fall	FSLEFRST	Subj. is standing, falls on the left side, and remains on the ground
	FSLEFRSTRC	Subj. is standing, falls on the left side, stays on the ground for a while and then gets up again
	FSLEFRWK	Subj. walks, falls on the left side, and remains on the ground
	FSLEFRWKRC	Subj. walks, falls on the left side, stays on the ground for a while and then gets up again.

(*continued*)

Table 1. (*continued*)

Type	Activity name	Description
Right side fall	FSRIFRST	Subj. is standing, falls on the right side, and remains on the ground
	FSRIFRSTRC	Subj. is standing, falls on the right side, stays on the ground for a while and then gets up again
	FSRIFRWK	Subj. walks, falls on the right side, and remains on the ground
	FSRIFRWKRC	Subj. walks, falls on the right side, stays on the ground for a while and then gets up again
Pick up object from floor with bending	APBEST	Subj. is standing, bends, picks up an object on the floor, and then stands up again
	APBEWK	Subj. walks, bends, picks up an object on the floor, and then stands up again
Pick up object from floor with squatting	APSQST	Subj. is standing, squats, picks up an object on the floor, and then stands up again
	APSQWK	Subj. walks, squats, picks up an object on the floor, and then stands up again
Sit and get up from the chair	ASCHST	Subj. is standing, sits on a chair, and then stands up again
	ASCHWK	Subj. walks, sits on a chair, and then stands up again
Sit and get up from the couch	ASSOST	Subj. is standing, sits on a couch, and then stands up again
	ASSOWK	Subj. walks, sits on a couch, and then stands up again

As a first possible discriminating feature, we considered the mean value of the speed of fall over all the test repetitions performed by each subject, and over all the subjects performing the same type of fall. The results obtained for the four classes considered are shown in Fig. 4. Apart from the case of FFOKCH-ST fall, that shows a quite high variability over the different 17 subjects who performed the test, with a much larger 95% confidence interval, it is not possible to say that the average speed of fall alone can be taken as a feature able to discriminate in a clear fashion the different types of falls simulated. In fact, both the mean values and the 95% confidence intervals are very similar over the four classes of falls analyzed.

We then move to analyze a different quantity, i.e. the mean difference between the ending and the starting values of the z coordinate, for each type of fall, which we call Δz. Figure 5 shows the mean value of Δz over all the test repetitions performed by each subject, and over all the subjects performing the same type of fall, again for the four classes considered before. In this case, it is quite evident that falls belonging to the class FBESFR may be grouped into a cluster (denoted as **A** in the graph) that is distinguishable from falls belonging to the classes FFOKCH, FFOKFR and FFOKSO, that are grouped into cluster **B**. Backward falls ending sitting in cluster **A** show a greater mean Δz than forward falls ending on the knees in **B**, irrespective of the different subjects' physique (the details regarding the subjects who performed the experiments are provided in Table 2). On the contrary, it is not possible to rely on the mean Δz values to discriminate among the different types of forward falls ending on the knees, that have been labeled as group **c**, **d**, and **e** in Fig. 5.

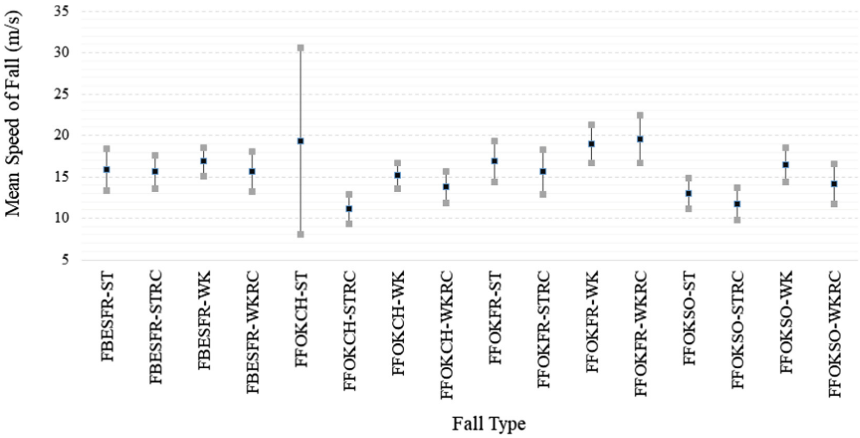


Fig. 4. Mean speed of fall along the z axis (black squares) and 95% confidence interval for the classes FBESFR, FFOKCH, FFOKFR and FFOKSO, averaged over all the subjects executing the test

Table 2. Information about the 17 voluntary users involved in the test phase.

Number	Gender	Age	Height (cm)	Weight (Kg)
1	Female	40	162	60
2	Female	29	170	74
3	Female	25	160	52
4	Male	30	176	65
5	Male	55	173	80
6	Male	21	169	58
7	Male	21	178	70
8	Male	23	175	59
9	Male	28	178	74
10	Male	28	160	76
11	Male	26	182	73
12	Male	40	187	87
13	Male	21	189	80
14	Male	22	167	64
15	Male	22	170	72
16	Male	21	188	78
17	Male	21	177	78
Mean		28	174	71
Std		9.3	9.2	9.5

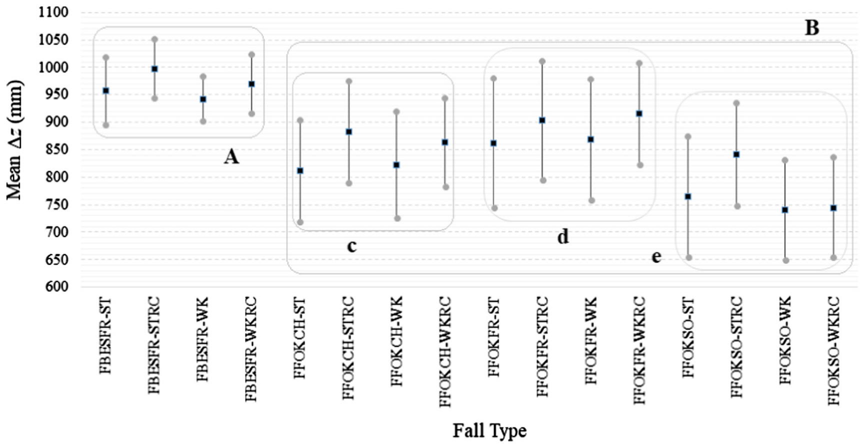


Fig. 5. Mean Δz values and 95% confidence intervals for the classes FBESFR, FFOKCH, FFOKFR and FFOKSO, averaged over all the subjects executing the test

4 Conclusion

In this paper we provided a preliminary investigation about the possibility to use the speed of fall as a feature to discriminate among different types of falls, obtained by processing the depth data provided by a Kinect device placed in top view with respect to the subject. This approach differs from those usually found in the literature, as the adoption of Kinect in top view makes not available the automatic detection of the subject's joints, which are output by the sensor when used in front view with respect to the monitored person. The speed of fall is obtained indirectly, by processing the depth frames through the identification of the person's blob in the image, and tracking of its relative distance from the sensor during the sequence of frames referred to the fall. The outcomes of this first attempt show that the average speed of fall alone is not enough to discriminate among falls, but this feature can be used in conjunction with others, like the variation of the vertical coordinate due to the fall, to improve the classification reliability. As already stated, this is just a preliminary study, but, based on the promising outcomes obtained, we aim to fully exploit the dataset collected through the experiments with 17 subjects, to improve our analysis and also include the evaluation of the activities of daily living.

Acknowledgment. This work was partly supported by Università Politecnica delle Marche under the project "PIMRA: Procedure Innovative per il Monitoraggio Remoto della Attività motoria e cardiovascolare". The authors would also like to acknowledge the contribution of the COST Action IC1303 AAPELE - Architectures, Algorithms and Platforms for Enhanced Living Environments.

References

1. Delahoz, Y.S., Labrador, M.A.: Survey on fall detection and fall prevention using wearable and external sensors. Sensors **14**(10), 19806–19842 (2014). https://doi.org/10.3390/s141019806. http://www.mdpi.com/1424-8220/14/10/19806
2. Gasparrini, S., Cippitelli, E., Spinsante, S., Gambi, E.: A depth-based fall detection system using a kinect sensor. Sensors **14**(2), 2756–2775 (2014). https://doi.org/10.3390/s140202756. http://www.mdpi.com/1424-8220/14/2/2756
3. Han, J., Shao, L., Xu, D., Shotton, J.: Enhanced computer vision with microsoft kinect sensor: a review. IEEE Trans. Cybern. **43**(5), 1318–1334 (2013). https://doi.org/10.1109/TCYB.2013.2265378
4. Spinsante, S., Cippitelli, E., De Santis, A., Gambi, E., Gasparrini, S., Montanini, L., Raffaeli, L.: Multimodal interaction in a elderly-friendly smart home: a case study. In: Agüero, R., Zinner, T., Goleva, R., Timm-Giel, A., Tran-Gia, P. (eds.) MONAMI 2014. LNICST, vol. 141, pp. 373–386. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16292-8_27
5. Spinsante, S., Gambi, E., De Santis, A., Montanini, L., Pelliccioni, G., Raffaeli, L., Rascioni, G.: Design and implementation of a smart home technological platform for the delivery of AAL services: from requirements to field experience, pp. 433–456. Healthcare Technologies. Institution of Engineering and Technology (2016). https://doi.org/10.1049/PBHE006E_ch21
6. Springer, S., Yogev Seligmann, G.: Validity of the kinect for gait assessment: a focused review. Sensors **16**(2), 1–13 (2016)
7. Wikipedia: Definition of sensor (2017). https://en.wikipedia.org/wiki/Sensor
8. Yang, L., Ren, Y., Zhang, W.: 3D depth image analysis for indoor fall detection of elderly people. Digit. Commun. Netw. **2**(1), 24–34 (2016). https://doi.org/10.1016/j.dcan.2015.12.001. http://www.sciencedirect.com/science/article/pii/S2352864815000681

How to Spread Kindness: Effects of Rewarding Elements Within a Persuasive Application to Foster Prosocial Behavior

Johanna Schäwel(✉) and Nicole C. Krämer

Social Psychology: Media and Communication, Department of Computer Science and Applied Cognitive Science, University Duisburg-Essen, Forsthausweg 2, 47057 Duisburg, Germany
{johanna.schaewel,nicole.kraemer}@uni-due.de

Abstract. For investigating whether rewarding elements within a persuasive web-application for increasing prosocial behavior are effective, we conducted an experimental field study with a self-developed web-app ($N = 42$). Two different versions of the persuasive web-app (high/low persuasive) to foster good deeds were implemented and examined during a three-week period. In both versions, the evaluation of the user-interface correlated positively with the execution of the target behavior, i.e. performing good deeds. Also, the availability of a user profile and a ranking was positively associated with performing prosocial behavior. However, there was neither a significant difference regarding the intensity of app usage nor concerning the number of performed good deeds between experimental groups (high/low persuasive). Therefore, the availability of more persuasive features does not necessarily entail more persuasion. In conclusion, we derive design suggestions for developing persuasive mobile applications which provide benefits for (pro-)social life through encouraging people to perform good deeds.

Keywords: Prosocial behavior · Good deeds · Persuasive technology Self-monitoring · Rewards

1 Introduction

The enormous potential of persuasive technologies reveals itself when considering the amount of application fields and their empirically documented success concerning persuasive processes. Persuasive technologies are used for purposes such as the elimination of phobias [7] or saving resources [20]. While these fields are well analyzed, there are only rare insights concerning the area of prosocial behavior and inter-human relationships. However, especially in the light of fast-paced and anonymous everyday lives, the realm of social life could benefit from persuasive technologies that sensitize people regarding interpersonal relationships and prosocial behavior. To this aim, we built a persuasive web-application that encourages prosocial behavior in terms of performing good deeds through persuasive and rewarding elements. We strive to understand whether providing persuasive mechanisms contributes to encouraging the target behavior of doing good deeds. Therefore, we examined correlations between the

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 163–172, 2018.
https://doi.org/10.1007/978-3-319-76111-4_17

usage of provided persuasive features within a persuasive app and resulting behavior. Also, we analyzed differences between the effects of two different versions of the app, which was either high persuasive (extended features) or low persuasive (basic features).

2 Persuasion and Prosocial Behavior

Persuasion. Persuasion is the process of empowering people to change an attitude or a behavior towards an object or topic [6]. Persuasive technologies are interactive computer-based systems, supporting individuals to implement an anticipated revision of behavior or attitude by using empowering features [17]. Following Fogg [13], modifying behaviors and attitudes through a persuasive system mainly depends on an individuals' motivation, (perceived) abilities, self-efficacy [1, 3, 25] and on an adequate trigger causing a target behavior at a specific point of time. The effectiveness of a persuasive system can further be expanded by supplying features of personalization, social exchange [16], and principles of perceptibleness and immediateness [21]. Within a process of persuasion, rewards serve as additional motivators, reinforcing a behavioral change [8, 10, 11]. Studies regarding user activity in social networks indicate that rewards and competitive elements such as bonus points and leader boards increase the intensity of users' activity and do positively affect the power of persuasion [12, 18, 23]. Furthermore, users like to apply leader boards as tools for presenting themselves in a positive light [26]. Self-monitoring features, such as personalized behavior diagrams, graphs or diaries, further support users of a persuasive system in noticing and changing their behavior effectively. Due to the process of introspection, people receive another point of view and thus can detect and change unwanted habits [4].

As it is for all technological devices, the effectiveness of persuasive systems is further related to its user-friendliness [22], provided support, and usage intensity [15].

Prosocial Behavior. Prosocial behavior covers a broad spectrum reaching from helping other people in specific situations to being kind in everyday life [24]. Prosocial behavior means to do something good, e.g. to help others. In some cases, individuals anticipate positive returns for themselves [2, 5]. The prosocial target behavior we are trying to foster in this work is characterized by performing and documenting good deeds, encouraged through a persuasive system.

3 A Persuasive Application to Foster Prosocial Behavior

Based on prior research demonstrating the power of persuasive applications and systems [18], we developed and implemented the *Good Deeds-App*. This app allows to document and (self-)monitor good deeds. We implemented two versions of the app. As shown in Fig. 1, one with basic features (low persuasive version) and one with extended features such as bonus-points, a diagram (both within the profile) and a ranking (high persuasive version).

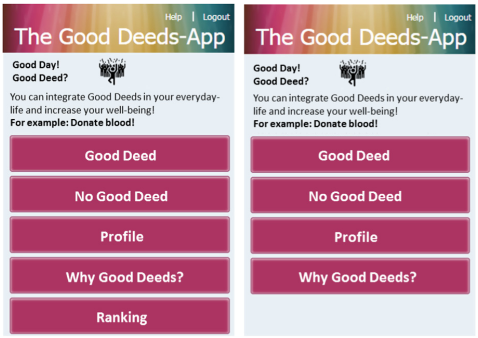


Fig. 1. Left: Extended features (high persuasive), right: Basic features (low persuasive).

3.1 Features of the Good Deeds-App

Starting page and main menu. The structure of the main menu is simple, allowing for easy usage even for unexperienced users. A randomized motivating welcome-text including an example for a good deed e.g. *"Sending a kind message to a family member or friend"* represented a behavioral trigger every time a user logged in.

"Good Deed". Text fields allow for documenting good deeds. The group with extended persuasive features was further able to collect bonus points, namely *Altruism-Stars* (five good deeds resulted into one *Altruism-Star*). Participants in this group were also informed about remaining good deeds for getting an *Altruism-Star*.

"No Good Deed". Self-reflection does not only comprise reflection of situations in which a target behavior was performed but also those in which it was not performed [13, 14]. Thus, users could click *"No Good Deed"* when they logged in but did not do a good deed. Subsequently they got a hint that it is not too late to do a good deed.

"Profile". In its basic version, the profile includes a diary displaying all performed good deeds in a chronological order and the number of good deeds per day. The app with extended features additionally provides the sub-menus *Altruism-Stars* and *Diagram* (Fig. 2). The diagram visualizes reported good deeds on a timeline, displaying an empty space when no good deed was performed.

"Why Good Deeds?". To strengthen awareness for social goods, a collection of facts explaining why doing good deeds is important and that it positively affects both, the executing and the receiving individual, was presented.

"Ranking". The ranking (only for high persuasive version), should initiate social comparison between users. The usernames and respective number of good deeds were visualized in a list that was available for every user within this group. The person at the leading position was additionally pushed via an encouraging emoticon.

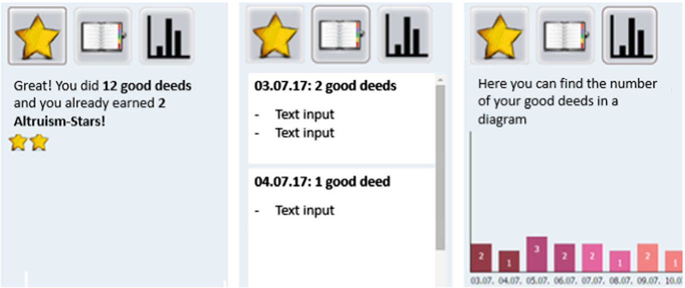


Fig. 2. Profile with extended features (left: *Altruism-Stars*, center: Diary, right: Diagram).

Trigger. Automated e-mails served as additional behavioral triggers for participants in both groups. If a participant did not log in to the app for a certain period he/she got a reminder to do so along with a suggestion to execute good deeds.

4 Derivation of Hypotheses

The aim of this empirical study is to investigate whether employing persuasive functions (e.g. self-monitoring, triggers) will empower users to conduct good deeds and whether more sophisticated functions (ranking, altruism-stars) will foster the target behavior even more strongly.

Since self-monitoring is a powerful persuasive tool [13, 14, 18, 20, 21, 26], we expect that monitoring one's own behavior through accessing the *Profile* within the app reinforces executing the target behavior of doing good deeds:

H1: *There is a positive correlation between the number of profile-accesses (for monitoring one's own behavior) and the number of good deeds. More precisely, the frequency of accessing the profile can significantly explain the number of good deeds.*

Further, a technology's interface can influence perceived credibility, usability and resulting joy of usage [19]. Hence, participants' evaluation of the user interface might positively correlate with utilizing the technology and consequently affect the number of good deeds:

H2: *There is a positive correlation between the evaluation of the persuasive app and the number of good deeds.*

Social comparison can increase behavioral engagement [4, 14, 18, 23]. Since one group had access to a ranking, we suppose that within this group the number of clicks on *Ranking* correlates with the number of good deeds:

H3: *There is a positive correlation between the number of clicks on the ranking and the total number of good deeds.*

Rewards arising from commendation, rankings, or bonus points can strengthen a person's willingness to change a habit [14, 23]. Concurrently, rewarding elements can increase positive experiences [10]. Based on this, we suggest that participants in the group with high persuasive features (extended version) use the app more intensely:

H4: *Participants who use the app with extended functions show a higher overall click frequency than participants who use the app with basic functions.*

H5: *Participants who use the app with extended functions show a higher click frequency for the Profile than participants who use the app with basic functions.*

Due to the different versions of the app (high/low persuasive), we suppose that there is a difference concerning the overall effect of using the introduced application. Consequently, we formulate the following hypothesis:

H6: *There is a difference between the groups regarding the number of good deeds.*

Additionally, we were interested in whether the different versions of the app fostered different types of good deeds. Following this question, we set the following explorative research question and qualitatively investigated users' text inputs:

RQ: *Is there a difference between the groups regarding the content of good deeds?*

5 Method and Sample

We conducted an experimental field study with a between subjects-design over a three-week period. The goals were to demonstrate that an app (operated under field conditions) can foster prosocial behavior and to analyze whether the employment of more persuasive strategies will increase this effect. Therefore, we varied the availability of rewarding elements (*Altruism-Stars*), positive feedback, a ranking and self-monitoring features (diary, diagram). Data was gathered via online questionnaires and behavioral log data of app usage.

In sum, 42 persons (23 female) participated in the study. In each condition, there were $n = 21$ participants (high persuasive: 8 female, $M_{age} = 28.05$, $SD_{age} = 8.45$ vs. low persuasive: 15 female, $M_{age} = 26.19$, $SD_{age} = 7.65$).

5.1 Variables

We investigated participants' satisfaction with the user-interface by means of adapted items from the *Questionnaire for User Interface Satisfaction* by Chin, Diehl and Norman [9], ($\alpha = .882$). For analyses, we used items concerning "overall impression" (e.g. "My impression of the app is 1 = horrible/10 = wonderful"), "screen" (e.g. "The arrangement of information on the screen is 1 = hard to read/10 = easy to read"), "terminology and system information" (e.g. "Terminology is 1 = not consistent/10 = consistent") and "learning" ("Learning how to use the app was 1 = hard/10 = easy") of the mentioned scale [9]. Furthermore, the frequency of clicks on every button as well as the number of active days within the app have been tracked. The number of active days within the web app was determined by summing up all days on which at least one menu item was clicked. To answer the stated research question, users' text inputs (good deeds) were qualitatively analyzed and clustered into categories by two independent coders ($\kappa = .922$, $p < .01$).

6 Results

6.1 Descriptive Results

We recommended to use the application at least once a day. However, in sum, participants spent between two and 16 days within the app (M = 6.00, SD = 2.83).

Good Deeds. Altogether, participants stored 234 good deeds (M = 5.57, SD = 2.96). Most commonly, people documented between three and six good deeds. On average, female participants stored more good deeds (M = 6.22, SD = 3.06) than male participants (M = 4.79, SD = 2.67).

Number of clicks. The overall number of clicks was 970 (M = 23.10, SD = 20.08). Participants in the group with extended functions had a total click rate of 520 (M = 24.77, SD = 27.04), participants in the group with basic functions clicked 450 times (M = 21.43, SD = 9.45).

6.2 Hypotheses Testing

H1. We conducted a regression analysis revealing a significant positive relation between the frequency of profile accesses and the number of good deeds (R^2 = .339, F(1,40) = 20.54, β = .582, t = 4.53, p < .001). The frequency of accessing the profile explained 34% of the variance of documented good deeds. Consequently, the first hypothesis was supported.

H2. To examine possible relations between the evaluation of the user-interface and the number of good deeds we conducted several analyses of correlation. There were significant positive correlations between two dimensions of the scale for evaluation of the user-interface and the number of good deeds, namely for "learning" (r = .352, p < .05) and "terminology and system information" (r = .356, p < .05). Therefore, hypothesis 2 can partially be accepted.

H3. The assumption of a positive relation between the frequency of accessing the ranking and the number of good deeds (high persuasive version) was tested via an analysis of regression and indicated a significant positive relation (R^2 = .249, F(1,19) = 6.31, β = .449, t = 2.51, p < .05). The frequency of clicking *Ranking* explained 25% of variance of the number of good deeds. We therefore accept hypothesis 3.

H4. We need to reject the fourth hypothesis since there was no statistically significant higher click rate for participants using the app with extended functions.

H5. Concerning the number of profile visits we found, contrary to the hypothesis, that participants using the app with basic functions accessed their profile more often than participants using the app with extended features (F(1,39) = 12.00, p < .001, η^2 = .235, user-interface evaluation was considered as a control variable). Therefore, hypothesis 5, in the way it was stated, needs to be rejected.

H6. Most central, we suggested a difference between the groups concerning the overall number of good deeds. The group with extended functions stored 100 good deeds (M = 4.76, SD = 2.32), whereas the group with basic functions in fact documented more good deeds (134, M = 6.38, SD = 3.34). However, the difference was not statistically significant. Therefore, hypothesis 6 needs to be rejected.

RQ. Analyses revealed seven categories of good deeds: "partner", "friends", "family", "strangers", "animals", "colleagues", "others". Most good deeds were performed for "strangers", whereby participants using the app with extended functions stored 37 (M = .37, SD = .48), and participants using the app with basic functions stored 44 good deeds (M = .33, SD = .47) in that category. Fewest good deeds were conducted for "animals". Within the group with extended functions, three good deeds were documented in that category (M = .03, SD = .17), the group with basic functions did five good deeds in that category (M = .04, SD = .19). Altogether, participants using the app with extended functions did significantly more good deeds within "partner" (χ^2 (1, N = 42) = 5.81, $p < .05$), while participants of the other group did significantly more good deeds within "family" (χ^2 (1, N = 42) = 5.29, $p < .05$).

7 Discussion

The experimental field study, which for the first time tested the effects of various persuasive strategies on performing prosocial behavior yielded numerous interesting results. Among others, we identified which features are most important in fostering prosocial behavior and further addressed the question whether more features lead to more and stronger effects.

We found that the amount of performed good deeds is attributable to the quantity of accessing the *Profile* (H1) and conclude that the functionality of an individual profile can help to perform a target behavior. In line with previous findings [13, 14, 18, 26] this indicates that self-monitoring is indispensable for a persuasive technology. Even if we acknowledge that the causality can also be reversed, meaning that participants, after storing a good deed, were interested in checking their profile, it is still remarkable that users felt the urge to check on their "status" instead of storing a good deed and closing the app. Since the profile differs between the groups with respect to provided self-monitoring functions, but the relation is present for the whole sample, it can be assumed that the functionality of the overall opportunity of self-observation was relevant for the process of persuasion, while the extended features of different self-monitoring functions do not seem to be particularly relevant.

Data further show that positive user-interface evaluations are related to the likelihood of storing good deeds (H2). This underlines the importance of the app's functionality and a positive evaluation by the users. In line with this, we want to acknowledge that the interface received altogether positive ratings.

When analyzing the relation between clicks on the ranking and the number of reported good deeds (H3), we found a significant result demonstrating that the more often participants checked the rating, the more good deeds they reported. Even though this might merely indicate that people who entered the app to store a good deed afterwards checked on their position in the ranking (instead of purposefully look this up), this still means that the ranking was relevant for them and thereby will have executed an effect. In line with prior findings [18] we conclude that striving for a good position in a ranking enhances the motivation for executing the target behavior. The ranking potentially served as a type of self-control and provided additional rewards when reaching a good position. However, it did not lead the participants in the group with extended functions to perform more good deeds than the group with basic functions. The persuasive power therefore seems to be restricted.

Examining hypothesis 4, we found that the groups did not significantly differ regarding overall click frequency. However, we found a significant difference when only investigating the frequencies of clicking the ranking. Contrary to our hypothesis (5), participants from the group with basic functions actually checked their *Profile* more often. It can also be derived that the power of persuasion does not increase proportionally with the number of provided self-monitoring-features but that one adequate measure for self-observation might be sufficient.

Contrary to our hypothesis we did not find a significant difference between the groups regarding the number of good deeds (H6). It is important to note that this is, for example, not merely due to lack of statistical power but that on a descriptive level the group with only basic functions even reported more good deeds than the group with extended persuasive features. It can be assumed that the participants of the experimental group with extended features were equally satisfied with available functions and provided rewards, such as the participants of the group with basic functions were pleased with the provided functions. Each group for itself may have perceived an adequate support. Therefore, the execution of good deeds was equally supported and encouraged. Strikingly, the most rewarding element for a conducted good deed, the *Altruism-Star*, was not sought-for. Only eleven participants documented five good deeds and were therefore rewarded with an *Altruism-Star*. In order to check whether the star – once received – will be able to increase the motivation to use the app and to perform more good deeds, future studies should consider to provide such bonuses earlier. Furthermore, usage phases should take a longer period of time.

Most good deeds were stored within the category “strangers” (RQ). This category is often neglected under conventional circumstances in terms of good deeds. The effectiveness of the *Good Deeds-App* has been confirmed exactly in this area. In addition, numerous good deeds within the categories “friends” and “colleagues” were accomplished, whereas the number of good deeds within the categories “partner” and “family” were lower. It can be assumed that performing good deeds for the immediate environment requires less support by a technology or that assistance and prosocial behavior within the immediate social environment are taken for granted and were therefore not documented.

8 Conclusion

The aim of this study was to investigate the influence of persuasive elements on the likelihood of doing good deeds encouraged through an application for enhancing prosocial behavior. We varied availability of rewarding elements such as bonus points

(*Altruism-Stars*) and positive feedback as well as a ranking-list and self-monitoring functions (diary, diagram), for testing if more features yield more target behavior.

The current study indicates that this is not the case: The availability of more features and functions did neither lead to more reported target behaviors nor to extended usage of the app and its features. Considering the data as a whole, we conclude that the most important functions were sufficiently represented in the basic version. Providing more persuasive opportunities does not necessarily help in terms of larger persuasive power. One of the functions that was not available in the basic version (ranking) yielded indeed enormous interest by the respective group and was also shown to be related to reporting the target behavior but was not able to spur behavior in a sufficient manner. The execution of the target behavior was shown to be related to usage of the profile as a self-monitoring feature. Moreover, we can derive that the positive evaluation of the user interface plays an important role for using the app as well as for the performance of the target behavior.

In spite of the mentioned limitations our work contributes to the state of the art by demonstrating that persuasive technologies can foster prosocial behavior and that more features do not necessarily lead to more persuasion especially if they address redundant functions and strategies. Nevertheless, persuasive technologies fostering prosocial behavior are promising for improving interpersonal relationships and well-being for a large number of people.

References

1. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. **50**, 179–211 (1991)
2. Aronson, E., Akert, R.M., Wilson, T.D.: Sozialpsychologie, 8th edn., p. 398. Pearson Deutschland GmbH, Hallbergmoos (2014)
3. Bandura, A.: Self-Efficacy: The Exercise of Control. Freeman, New York (1997)
4. Bang, M., Torstensson, C., Katzeff, C.: The powerhouse: a persuasive computer game designed to raise awareness of domestic energy consumption. In: Ijsselsteijn, W., de Kort, Y., Midden, C., Eggen, B., van den Hoven, E. (eds.) Persuasive Technology, pp. 123–132. Springer Verlag, Berlin (2006). https://doi.org/10.1007/11755494_18
5. Batson, C.D.: The Altruism Question: Toward a Social-psychological Answer, p. 50. Lawrence Erlbaum Assoc Inc., Hillsdale, New Jersey (1991)
6. Bohner, G., Dickel, N.: Attitudes and attitude change. Ann. Rev. Psychol. **62**, 391–417 (2011)
7. Botella, C., García Palacios, A., Rivera, B., María, R., Quero Castellano, S.: Cybertherapy: advantages, limitations, and ethical issues. PsychNology J. **7**, 77–100 (2009)
8. Busch, M., Schrammel, J., Tscheligi, M.: Personalized persuasive technology - development and validation of scales for measuring persuadability. In: Berkovsky, S., Freyne, J. (eds.) Persuasive Technology, pp. 33–38. Springer, Berlin (2013). https://doi.org/10.1007/978-3-642-37157-8_6
9. Chin, J.P., Diehl, V.A., Norman, K.L.: Development of an instrument measuring user satisfaction of the human-computer interface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 213–218. ACM, New York (1988)

10. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9–15. ACM, New York (2011)
11. Deterding, S.: Eudaimonic design, or: six invitations to rethink gamification. In: Fizek, S., Fuchs, M., Ruffino, P., Schrape, N. (eds.), pp. 305–331. Meson Press, Lüneburg (2014)
12. Farzan, R., DiMicco, J.M., Millen, D.R., Dugan, C., Geyer, W., Brownholtz, E.A.: Results from deploying a participation incentive mechanism within the enterprise. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 563–572. ACM, New York (2008)
13. Fogg, B.J.: A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology, pp. 40:1–40:7. ACM, New York (2009)
14. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann, Elsevier, San Francisco (2003)
15. Fogg, B.J., Tseng, H.: The elements of computer credibility. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 80–87. ACM, New York (1999)
16. Fogg, B.J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., Treinen, M.: What makes web sites credible?: a report on a large quantitative study. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 61–68. ACM, New York (2001)
17. Fogg, B.J., Cuellar, G., Danielson, D.: Motivating, influencing, and persuading users. In: Sears, A., Jacko, J.A. (eds.) The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, 2nd edn, pp. 133–147. Taylor & Francis Group, New York (2007)
18. Foster, D., Lawson, S., Blythe, M., Cairns, P.: Wattsup?: motivating reductions in domestic energy consumption using social networks. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, pp. 178–187. ACM, New York (2010)
19. Halko, S., Kientz, J.A.: Personality and persuasive technology: an exploratory study on health-promoting mobile applications. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) Persuasive Technology, pp. 150–161. Springer, Berlin (2010). https://doi.org/10.1007/978-3-642-13226-1_16
20. Kappel, K., Grechenig, T.: "Show-me": water consumption at a glance to promote water conservation in the shower. In: Proceedings of the 4th International Conference on Persuasive Technology, pp. 26:1–26:6. ACM, New York (2009)
21. Laschke, M., Diefenbach, S., Heidecker, S., Hassenzahl. M.: Transformationale Produkte: Acht Konzepte zum schonenden Umgang mit Ressourcen. In: Ziegler, J., Schmidt, A. (eds.) Mensch und Computer 2010: Interaktive Kulturen, pp. 189–194. Oldenbourg Verlag, München (2010)
22. Lee, Y., Kozar, K.A.: Understanding of website usability: specifying and measuring constructs and their relationships. Decis. Support Syst. **52**, 450–463 (2012)
23. Mekler, E.D., Brühlmann, F., Opwis, K., Tuch, A.N.: Do points, levels and leaderboards harm intrinsic motivation?: An empirical analysis of common gamification elements. In: Proceedings of the First International Conference on Gameful Design, Research, and Applications, pp. 66–73. ACM, New York (2013)
24. Schwartz, S.H., Howard, J.A.: In: Bierhoff, H.-W. (ed.) Sozialpsychologie: Ein Lehrbuch, pp. 96–139. Kohlhammer, Stuttgart (1981)
25. Spahn, A.: And lead us (not) into persuasion…? Persuasive technology and the ethics of communication. Sci. Eng. Ethics. **18**, 633–650 (2012)
26. Sra, M., Schmandt, C.: Spotz: a location-based approach to self-awareness. In: Berkovsky, S., Freyne, J. (eds.) PERSUASIVE 2013. LNCS, vol. 7822, pp. 216–221. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-37157-8_26

Theorizing Gamified Virtual Reality Approach to Overcome Fear of Height

Imran Khaliq(✉), Jc Fowles, and Callan Moore

Media Design School, 1010 Auckland, New Zealand
imran.khaliq@mediadesignschool.com,
{jc.fowles,callan.moore}@mediadesign.school.nz

Abstract. The use of virtual reality as a form of exposure therapy for people who suffer from acrophobia (fear of height) has been proved and tested by multiple studies. In this paper, we initiate gamified virtual reality approach to overcome fear of height by providing a design implementation theory. We call this theory High Engagement and Low Intensity-Low Engagement and High Intensity. This theory adds a gamified element to the virtual environment. The idea here is to have game engagement at its highest and intensity of the acrophobia at its minimum in the virtual environment at the start of the treatment. As the treatment progresses, the engagement in the game starts decreasing and the intensity of the phobia starts increasing. The implementation provides two parameters (engagement and intensity) that could be adjusted independently according to the patient needs. We also evaluate the base of our theory through a pilot study.

Keywords: Acrophobia · Fear of height · Gamification · Treatment
Virtual reality

1 Introduction

Acrophobia, or more commonly known as the fear of heights, is a debilitating anxiety order that many people suffer from. Between two to five percent of the world population is afflicted by this mental disorder with twice as many females as males affected [1]. Acrophobic people are afraid to be in high places, though "high" is a relative term. Some people are fearful of being on a three story balcony with no railings. However, in the more extreme cases, simple standing on a single chair on the ground is high enough to bring about a phobic reaction. According to [2], the fear of heights would need to be overpowering and inhibiting to term as a phobia. In [3], Stanley Rachman created a model called the *pathways of fear* where he has identified three ways fear acquisition. They are (1) Learning by direct experience, (2) Learning by observing others, and (3) Learning through information/instruction. In addition to the Stanley Rachman model, in [4], fourth pathway called *non-associated perspective* is identified that cause fear. Acrophobia, like many other phobias, has symptoms that relate to

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 173–182, 2018.
https://doi.org/10.1007/978-3-319-76111-4_18

panic and anxiety attacks. These symptoms are often followed by other more physical signs including sweating, abnormal breathing, accelerated heartbeat, trembling, disorientation, nausea, and even dizziness [5]. In acute cases, sufferers can become so overwhelmed with their fear that their bodies seize up and they are unable to function or move [1]. The author of the book, "Diseases of the Human Body", mentions that the phobics are often aware that their fear is irrational. However even knowing this, they are unable to confront or control their fears [6]. Estimates from the study [7] on refusal and acceptance of treatments for phobias showed that 60–80% of people with specific phobias are reluctant to pursue help or treatment. Acrophobics can be coached to develop mechanisms for dealing with their fear through the help of therapists. Therapists use different types of behaviour therapies to achieve results. Such therapies are build upon the assumption that the sufferer has learned this response to being in particular environments or situation and hence can be treated as if one can learn a reaction then it can be unlearn too. One such therapy is called cognitive-behaviour therapy (CBT). The CBT involves introducing a patient to their fears while trying to inspire them to alter their thought processes about their fear. A common process of CBT is *system desensitization* in which a patient is gradually exposed to the stimuli. This treatment is administered in two stages, the first relaxation training and imagined situations going from least to most fearful. The second stage is actual exposure to the feared situation from the least threatening to the most. Though it can take a long time for these treatments to be effective, taking months or even years in some cases. Some medications, such as tranquilizers, anti-depressants, anti-anxiety, beta blockers and sedatives, can also be utilised in the treatment of acrophobia. These medications will not cure the condition by themselves however. They are used to help minimise or relieve some of the symptoms of acrophobia, especially anxiety, in order to allow patients to cope better.

1.1 VR Exposure Therapy

There have been multiple studies concerning treatment of phobias using Virtual Reality (VR). The first proper study was done by Carlin et al. in [8]. Research projects around the world were sparked by the encouraging results of this venture such as fear of public speaking [9], fear of driving and claustrophobia [10], agoraphobia [11]. VR treatment works by creating a computer generated simulation that would mimic an in-vivo situation where the user is exposed to their fear. The VR has advantages over in-vivo such as VR is less expensive and less time consuming; VR provides safer environment; VR removes the risk of any possible public humiliation for the patient. Taking into account these various studies [12–15] done on using VR in acrophobia treatment and their results, there is now a considerable amount of evidence to support the use of VR exposure as an effective means of treating acrophobia.

2 Gamification Theory

Our purpose is to take the idea of using VR in the treatment of acrophobia and gamify it. The term *Gamification* is an umbrella term used when video game elements are added to a non-gaming system [16]. The purpose of gamifying a non-gaming system is to motivate and engage the participants. Our main purpose of gamifying virtual reality treatment for acrophobia is to draw patient's attention away from thoughts associated with heights. This is achieved by engaging the patient in playing a virtual reality game. As many studies have proved that a person can be distracted from pain and the associated fear by involving in some fun activity. This distraction is effectively achieved through the use of virtual reality environment as proved by various studies [17–20]. We propose a design implementation theory of how to design the gamified virtual reality environment for treating acrophobia. We call this theory High Engagement and Low Intensity - Low Engagement and High Intensity. In short, we use HELI-LEHI for High Engagement and Low Intensity - Low Engagement and High Intensity. The HELI-LEHI theory is twofold: first is the high to low game engagement and second, low to high phobia intensity. The high to low gameplay engagement idea is to have the engagement at its highest at the start of the treatment. As the treatment progresses, the game engagement decreases to the point that it is very minimal (or non-existent). The purpose of engaging a participant in the game is to distract the participant from her fear while she is still subconsciously going through it. The more engaging the game is, the less the participant will notice the height stimuli at the time. Though, this may sound counterproductive as the point of the virtual reality treatment is exposure. It is not, because as the treatment progresses the engagement is lessened to the point where it is almost non-existent. This leaves the participant to be fully exposed to the height without any distraction. It is important to remember not to distract the participant entirely to the extent that these stimuli no longer induce the phobia. The participant needs to be exposed to the height stimuli and know about it, however, not primarily focused on it. This distraction goes hand in hand with the idea of low phobia intensity to high phobia intensity part of the theory. We suggest that participant is to be exposed to the least fear evoking situation and gradually moved to a more serious situations as the treatment progresses - following similar approach to systematic desensitization in [21]. In our case the phobia is about height, so the participant can start close to ground level and as the treatment progresses the participant moves to higher and higher environments. The HELI-LEHI theory introduces participants to their phobia more gradually using the two parameters - engagement and intensity, allowing the participants to conquer their phobia one step at a time (Fig. 1).

Our HELI-LEHI theory has three benefits: (1) the game element adds fun to the treatment procedure which provides motivation to the patient who now becomes a player. (2) the game engagement distracts patient's mind temporarily from not thinking about the height. This way it provides extra level of ease, for the benefit of the patient, in the treatment. (3) it provides two independent

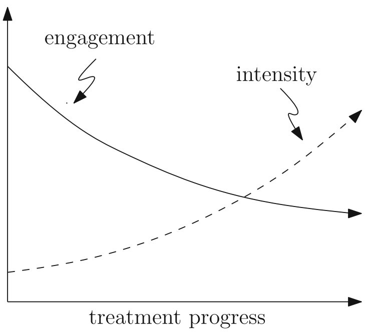


Fig. 1. HELI-LEHI

parameters, engagement and intensity, that makes the treatment more adjustable to the needs of each individual patient.

3 Virtual Reality Game

For the purpose of this study - the idea of "crossing a bridge" and "dodging incoming projectiles" at different floors was chosen. The reason for this was largely because the idea is easily scalable. Crossing a bridge allows both the intensity and engagement to be scaled up or down mutually exclusive of each other. The intensity of the phobic stimuli is generated via the current height that the player (the patient is now becomes a player) has achieved. Engagement is created by the projectile system which has multiple variables (projectile speed, fire rate, amount of projectiles) to increase or decrease how much attention is required by the player. Re-playability is also a factor as this experience is designed for players to partake in multiple times. Puzzles once solved do not offer the same interaction in a subsequent play through.

The gamified experience that is created for the experiment starts at ground floor. Movement is controlled with an Xbox 360 controller. The viewport is created with an Oculus Rift Head Mounted Display (HMD). The player begins on a sidewalk of a city street surrounded in all directions by tall high rise buildings and cars driving along the road. The player is not permitted to roam the streets and is limited to the sidewalk in front of the main building only. This ground floor stage has no real interaction with either the fear itself or the gamified elements of the simulation - it is more of an adjustment area for the player to enter and get acquainted with the simulated world. Background sounds are played that emulate basic city streets including car noises and also player's footsteps sounds. These sounds are played at ground floor so, the higher the player goes in the simulation the quieter these noises become. Once, the player has adjusted to the simulated world and are ready to proceed, there is an elevator waiting on the sidewalk ready to take them to the first real part of the simulation. Entering the elevator automatically triggers the doors to close and the elevator starts its slow ascent. The elevator has two glass windowed sides, so that the player is gradually introduced to the height as they can see the movement occurring. This version of the virtual reality game contains three levels only. It is meant to serve as the prototype.

3.1 Game Levels

When the elevator reaches the third floor of the main building the elevator stops and the player can move out onto a balcony. The balcony allows the player to move around and start to feel the sensation of height. There is an opening in the middle with a small platform that the player can stand on. This is the first part of the gamified experience. The small platform is part of the bridge and will start moving across a wire to the other side of the road to another building. At this stage the player can actually walk off the platform completely. This is intentionally allowed so that the player does not fall prey to knowing that they cannot. However, the aim of this experiment is not to terrify the player, if the player character (not the player herself) starts to fall, the simulation fades to black very quickly and then the game needs to be restart. When the player first steps onto the platform a door from the opposite building opens up and an installation of three projectile launchers are shown. On a short interval, two of the projectile launchers will fire creating a gap that the player needs to maneuver their 'head' into, in order to dodge the projectiles (blue in color) – as seen in Fig. 2. Oculus Rift head motion tracking is used in order to achieve this. The Oculus HMD allows the camera to be moved relative to the player, creating a balancing act for the player (the projectiles are rigged to collide with the camera position only). This movement of the head affects the camera only and causes no movement so the player cannot trigger a fall in the simulated world via this. Once the player crossed the bridge, the door with the launcher installation closes and the launchers stop firing. Stepping off the platform puts the player on another balcony with another glass elevator to take them to the next level.

The setup of this second bridge is almost identical to the first one. The interaction takes place on the seventh floor of the building raising the intensity of the acrophobia stimuli. In order to bring more attention to the height factor, the game is lessened so that the player is less engaged. The launcher installation is triggered in the same way except the fire rate is decreased so there is less projectiles to dodge. The launcher placement also only selects one launcher to fire rather than two making the dodging easier and less attention consuming. This is the intermediate stage of HELI-LEHI - medium intensity and medium engagement. The player gets time in between the projectiles to notice their height and take in their surroundings. Again, there is an elevator available to transport the player to the next level. For this study, the next level is the final level.

In the final level, the engagement is at its minimum and intensity is at its highest point. We removed the game engagement at all so that to expose the player to the maximum intensity of the phobia. There is no projectiles to dodge, and the bridge is no longer a platform that automatically moves when the player is on it, this can be seen in Fig. 3. A glass bridge extends across the gap allowing the player to move across it at their own pace and really be able to see to interact with the high intensity of the height. The player upon reaching the opposing side is greeted by a door opening up with a victory message displaying that the simulation is over. This is the Low Engagement, High Intensity (LEHI) stage of the proposed theory. The assumption here is that the player is now at her final stage of the treatment and ready to be exposed to the maximum intensity of the height.

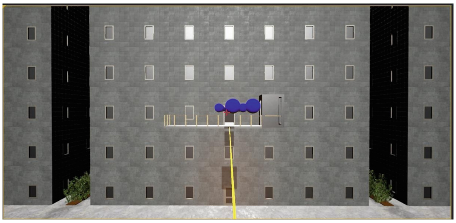


Fig. 2. First Bridge (Color figure online)

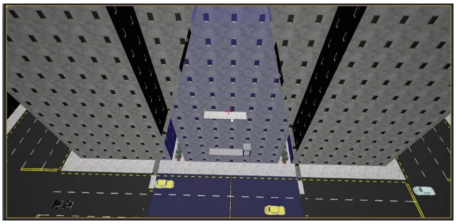


Fig. 3. Final Bridge

4 Method

In order to start the process of testing the HELI-LEHI theory, a small pilot study is the first logical step. There are two goals that we want to achieve in the pilot study. The first goal is to create a virtual environment that stimulates a sense of height. The second goal is to prove that the participants awareness level about their surroundings will decrease with the increase of the game engagement. The HELI-LEHI theory mainly bases on the second goal. Here, it should be noted that this is not the purpose of the pilot study to test how effective the gamified virtual reality approach is in the treatment of fear of height. We will conduct such experiment on acrophobia patients under the supervision of a qualified medical professional in a clinical environment in future. For the pilot study, we conducted two experiments to test the two goals. The results and related analysis are given below. A total of 15 volunteers, 4 females and 11 males aged between 18 to 38, were recruited from the Media Design School. Most of them were students. For our experiments, the participants are not required to be acrophobia patients.

4.1 Results and Analysis

Experiment 1: The purpose of the experiment is to make sure that our developed virtual world does stimulate a sense of height. To begin the experiment, we let the participants one by one to have a walk-through of all three bridges. In this experiment, the participants are not required to play the game (dodging projectiles) during crossing the bridges. The first bridge is at 3rd floor, the second bridge is at 7th floor, and the third bridge is at 12th floor. The participants started their journeys by entering the building, then went up to the third floor through elevator and crossed the first bridge. Similarly, the participants experienced the other two bridges. After finishing the final bridge, the participants came straight back to the ground floor, through elevator, to exit the building. The walk-through finished by exiting the building and hence, the experiment ended. Once the experiment became over, we recorded their responses about the following two statements: (1) I felt a sense of high up being in the virtual world, (2) I felt fear of height being in the virtual world. For each statement, we provided a 5-point scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree) to record their responses. The participants' responses about the first statement are provided in the Fig. 4. The participants' responses about the second statement are provided in the Fig. 5.

Experiment 1 Results Analysis: In Fig. 4, we can see that 13 out of 15 participants felt high up in the virtual world. It shows that the developed virtual world is good enough to carry on with. Figure 5 captures participants' fear of height. In Fig. 5, it is shown that, 11 out of 15 participants have felt fear of height when they were in the virtual environment. From the results, it is evident that the developed virtual world is effective and hence, our first goal is successfully achieved.

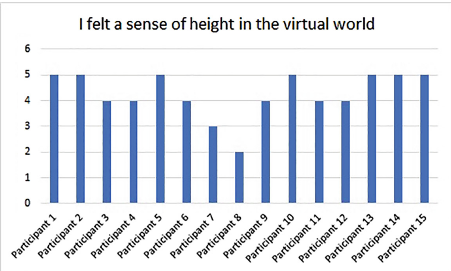


Fig. 4. Participants' response to sense of height in the VR world

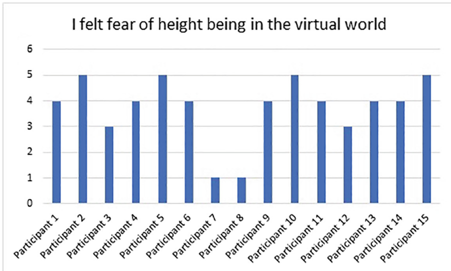


Fig. 5. Participants' response to fear of height in the VR world

Experiment 2: In this experiment, we want to test the awareness level of our participants while playing the game. Our claim is that the more engaging the game is the less the participants will be aware of their surroundings. We chose first and second bridge to carry out the experiment. We set, at first bridge, the game engagement at its highest and at final bridge, we reduced the game engagement drastically. We want to have a clear difference between game engagements at the two bridges. The participants first brought up to the first bridge, and let the participants to start playing the game. Once the participant played the game for a while, then we started moving the virtual platform upon which the participant is standing in the virtual world. As we started moving the virtual platform, at the same time we started recording time using a digital timer. We stop the timer as the player shouts to indicate us that he/she realized the movement of the virtual platform. We then move the participant to the second bridge to perform the same experiment. The participants' time at first bridge and final bridge are provided in the Fig. 6. The time of the participants recorded at the first bridge are represented in blue in colour and the time recorded at the second bridge are represented in orange in colour in the Fig. 6. It should be noted that for the sake of drawing the graph, we allocated 5 points to the participants who did not notice the movement of the virtual platform at all and crossed the bridge. In the Fig. 6, participants 1 and 7 failed to notice the movement during the entire time and hence 5 points were allocated. Moreover, we allocate 0 to the participants who noticed the movement of the virtual platform immediately or took less than a second.

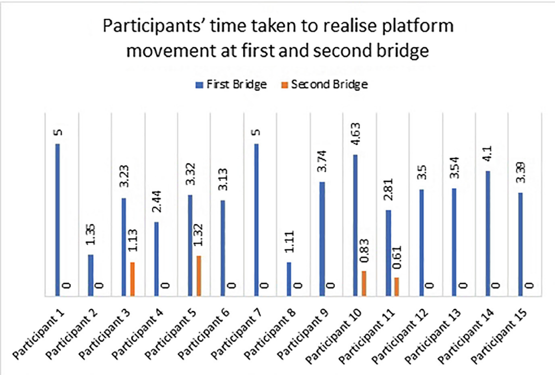


Fig. 6. Participants' level of awareness in the VR world (Color figure online)

Experiment 2 Results Analysis: At first bridge, when the engagement of the game was set at its highest, none of the participants realized the movement of the virtual platform immediately. All participants took more than 1 s to realize the movement of the virtual platform. Two of the participants (1 and 7) were fully immersed in the game during crossing the bridge and did not notice the movement of the virtual platform at all. At second bridge, when the engagement of the game was reduced drastically, 11 out of 15 participants were immediately noticed the movement of the virtual platform. The remaining 4 participants took time at most a little more than a second (1.32 s) to realize the movement of the virtual platform. Hence, we concluded that the more the participants were engaged in the game, the longer they took to notice the movement of the virtual platform, and hence, the less time their mind spent thinking about the height.

5 Conclusion

We initiated gamified virtual reality approach to overcome fear of height by providing a design implementation theory. We called this theory High Engagement and Low Intensity-Low Engagement and High Intensity. This theory provides two parameters independent of each other, intensity and engagement, to a medical professional. We implement the intensity parameter through bridges at different floors. The higher the bridge is, the more intense the fear of height will be. The engagement parameter is provided through the game (dodging projectiles). The engagement parameter is additional parameter that our theory provides. This engagement parameter is what makes our gamified virtual therapy differentiated from other virtual reality exposure therapies. The benefits of our gamified virtual reality approach is three folds. One, the game element adds fun to the treatment procedure which provides motivation to the patient who now becomes a player. Second, the game engagement distracts patient's mind temporarily from not thinking about the height. This way it provides extra level of ease, for the benefit of the patient, in the treatment. Third, having two independent parameters (engagement and intensity) in the hand of a medical professional makes the treatment more adjustable to the needs of each individual patient. We also evaluated the base of our theory by conducting a pilot study. The results of the pilot study mentioned above are in favour of our theory. For future work, we will test our theory by conducting a full user study in a clinical environment under the supervision of a qualified medical professional.

References

1. 11 curious Acrophobia statistics. In: Psychological Articles and Infographics, HRFnd (2014). (Online). http://healthresearchfunding.org/acrophobia-statistics/. Accessed 11 May 2016
2. Antony, M.M., Rowa, K.: Overcoming Fear of Heights: How to Conquer Acrophobia and Live a Life Without Limits. New Harbinger Publications, Oakland (2007)
3. Rachman, S.: The conditioning theory of fearacquisition: a critical examination. Behav. Res. Ther. **15**(5), 375–387 (1977)
4. Menzies, R.G., Clarke, J.C.: The etiology of phobias: a nonassociative account. Clin. Psychol. Rev. **15**(1), 23–48 (1995)
5. Maharjan, R.: Acrophobia: fear of height - causes, symptoms and treatment. In: Healthtopia 2016 (2016). (Online). https://www.healthtopia.net/disease/mental-health/phobia/acrophobia-causes-symptoms-treatment. Accessed 11 May 2016
6. Tamparo, C.D., Lewis, M.A.: Diseases of the Human Body, 5th edn, p. 134. F. A. Davis Company, Philadelphia (2011)
7. Garcia-Palacios, A., Botella, C., Hoffman, H., Fabregat, S.: Comparing acceptance and refusal rates of virtual reality exposure vs. In vivo exposure by patients with specific phobias. CyberPsychol. Behav. **10**(5), 722–724 (2007)
8. Carlin, A.S., Hoffman, H.G., Weghorst, S.: Virtual reality and tactile augmentation in the treatment of spider phobia: a case report. Behav. Res. Ther. **35**(2), 153–158 (1997)
9. Strickland, D., Hodges, L., North, M., Weghorst, S.: Overcoming phobias by virtual exposure. Commun. ACM **40**(8), 34–39 (1997)
10. Botella, C., Baños, R.M., Perpi, C., Villa, H., Alcaiz, M., Rey, A.: Virtual reality treatment of claustrophobia: a case report. Behav. Res. Ther. **36**(2), 239–246 (1998)

11. Malbos, E., Rapee, R.M., Kavakli, M.: A controlled study of agoraphobia and the independent effect of virtual reality exposure therapy. Australian New Zealand J. Psychiatry **47**(2), 160–168 (2013)
12. Rothbaum, B.O., Hodges, L.F., Kooper, R., Opdyke, D., Williford, J.S., North, M.: Virtual reality graded exposure in the treatment of acrophobia: a case report. Behav. Ther. **26**(3), 547–554 (1995)
13. Emmelkamp, P.M.G., Krijn, M., Hulsbosch, A.M., de Vries, S., Schuemie, M.J., van der Mast, C.A.P.G.: Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia. Behav. Res. Ther. **40**(5), 509–516 (2002)
14. Choi, Y.H., Jang, D.P., Ku, J.H., Shin, M.B., Kim, S.I.: Short-term treatment of Acrophobia with virtual reality therapy (VRT): a case report. CyberPsychol. Behav. **4**(3), 349–354 (2001)
15. Coelho, C.M., Santos, J.A., Silvrio, J., Silva, C.F.: Virtual reality and Acrophobia: one-year follow-up and case study. CyberPsychol. Behav. **9**(3), 336–341 (2006)
16. Deterding, S., Sicart, M., Nacke, L., OHara, K., Dixon, D.: Gamification. Using game-design elements in non-gaming contexts. In: Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA 2011 (2011)
17. Gold, J.I., Kim, S.H., Kant, A.J., Joseph, M.H., Rizzo, A.S.: Effectiveness of virtual reality for pediatric pain distraction during IV placement. CyberPsychol. Behav. **9**(2), 207–212 (2006)
18. Das, D.A., Grimmer, K.A., Sparon, A.L., McRae, S.E., Thomas, B.H.: The efficacy of playing a virtual reality game in modulating pain for children with acute burn injuries: a randomized controlled trial. BMC Pediatr. **5**, 1–10 (2005)
19. Gershon, J., Zimand, E., Lemos, R., Rothbaum, B.O., Hodges, L.: Use of virtual reality as a distractor for painful procedures in a patient with padiatric cancer: a case study. CyberPsychol. Behav. **6**, 657–661 (2003)
20. Kevin, M.M., Leonard, S.M.: The effectiveness of virtual reality distraction of pain reduction: a systematic review. Clin. Psychol. Rev. **30**, 1011–1018 (2010)
21. Wolpe, J.: Psychotherapy by reciprocal inhibition. Cond. Reflex: Pavlovian J. Res. Ther. **3**(4), 234–240 (1968)

The Fractal Dimension of Music: Geography, Popularity and Sentiment Analysis

Laura Pollacci[1(✉)], Riccardo Guidotti[2], Giulio Rossetti[1], Fosca Giannotti[2], and Dino Pedreschi[1]

[1] Department of Computer Science, University of Pisa, Pisa, Italy
{laura.pollacci,giulio.rossetti}@di.unipi.it, dino.pedreschi@unipi.it
[2] ISTI-CNR, Pisa, Italy
{riccardo.guidotti,fosca.giannotti}@isti.cnr.it

Abstract. Nowadays there is a growing standardization of musical contents. Our finding comes out from a cross-service multi-level dataset analysis where we study how geography affects the music production. The investigation presented in this paper highlights the existence of a "*fractal*" musical structure that relates the technical characteristics of the music produced at regional, national and world level. Moreover, a similar structure emerges also when we analyze the musicians' popularity and the polarity of their songs defined as the mood that they are able to convey. Furthermore, the clusters identified are markedly distinct one from another with respect to popularity and sentiment.

Keywords: Music data analytics · Hierarchical clustering
Sentiment pattern discovery · Multi-source analytics

1 Introduction

Music has been part of human civilization for centuries: it has been referred to as the *universal language*. Certainly, every culture has given birth to its own music "style": however, as time goes by, the constant collapse of physical barriers as well as the progressive reduction of geographical distances eased by the media and the world wide web caused an overall globalization of music.

During the last decade, the constant growing of on-line streaming services (such as Spotify, Apple Music, Last.fm) has made available to the public the widest choice of music ever. Emerging bands as well as famous ones can obtain a global visibility that was unimaginable only a few years ago. In this rapidly evolving scenario, music seems to have lost its geographical-cultural connotation: is it true that we are observing a growing standardization of music contents? Are there peculiar characteristics able to discriminate the music produced in a given region from the one produced in the country that contains it?

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 183–194, 2018.
https://doi.org/10.1007/978-3-319-76111-4_19

In this paper, leveraging a cross-service multi-level dataset, we study how geography affects the music production. Moreover, we study how artists producing a specific type of music are able to reach high popularity and how such popularity is often not related to a specific genre. Our data-driven investigation highlights the existence of a "*fractal*" structure that relates the technical characteristics of the music produced at regional, national and world level. Our structure reflects the well-knowledge characteristics of fractals objects theorized in mathematics since it exhibits a detailed pattern that repeats itself at increasingly big scales. Starting from emergent groups of an Italian region (Tuscany), moving to affirmed Italian artists and finally to a set of world-famous musician, we identify a multi-level set of profiles transversal to the classical concept of genre: we show how such profiles remains stable across the different geographical layers analyzed. Finally, we observed how the mood expressed by artists' songs as well as their popularity vary w.r.t. the multi-level traversal profiles they belong to. Once again, we observed the presence of the same structure that clearly emerges both at different geographical levels and over a multi-level set of profiles.

The rest of the paper is organized as follows. In Sect. 2 is discussed the literature involving music data analysis. Section 3 introduces the datasets used in the current study, and the preprocessing steps performed on them. In Sect. 4 we show and discuss the fractal structures emerging from the datasets and the music profiles that we were able to identify applying an unsupervised learning strategy. Section 5 highlights the relationships of the artists' popularity and music genres with respect to the groups discovered. Following the same approach, in Sect. 6 we analyze the relationships between the song lyrics and the music profiles enriched with sentiment analysis. Finally, Sect. 7 concludes the paper and discusses some future research directions.

2 Related Work

During the last decade, the music world has started receiving increasing attention from the scientific community. Several works [17,18] have analyzed data regarding on-line listening in order to model diffusion of new music genres/artists, as well as to analyze the behaviors and tastes of users.

Moreover, nowadays the on-line platform *Last.Fm* is offering a privileged playground to study different phenomena related to the on-line music consumption. In [19] was proposed a music recommendation algorithm by using multiple social media information and music acoustic-based contents. Hypergraphs developed on Last.fm model the objects and relations, and music recommendation is considered as a ranking problem. In [20] the authors studied the topology of the Last.Fm social graph asking for similarities in taste as well as on demographic attributes and local network structure. Their results suggest that users connect to "on-line" friends, but also indicate the presence of strong "real-life" friendship ties identifiable by the multiple co-attendance of the same concerts. The authors of [21] measured different dimensions of social prominence on a social graph built upon 70k Last.Fm users whose listening were observed for 2 years. In [22] was analyzed the cross-cultural gender differences in the adoption and usage of Last.Fm. Using data from Last.Fm and Twitter the authors of [23], designed a measure that describes diversity of musical tastes and explored its relationship with variables that capture socioeconomic, demographics, and traits such as openness and degree of interest in music. In [24] is shown how to define statistical models to describe patterns of song listening in an on-line music community. Finally, yet using Last.fm data, in [16] the authors defined a Personal Listening Data Model able to capture musical preferences through indicators and patterns, and showed how the employment of such model can provide to the individual users higher levels of self-awareness. They also discovered that all the users are characterized by a limited set of musical preferences, but not by a unique predilection.

In this paper, we are looking for dependencies between different levels of musical artist by developing a musical profile through clustering techniques. Also in [25–27] are used clustering techniques on musical data with different purposes. In [25], is studied the problem of identifying similar artists by integrating features from diverse information sources. In [27] a compression distance has been used to generate the similarities among music files and the paper shows some experimental results using these representations and compares the behavior of various clustering methods. Also in [26] is studied the problem of building clusters of music tracks in a collection of popular music in the presence of constraints and is presented an approach incorporating the constraints for grouping music by similar artists.

To the best of our knowledge, this work is the first attempt in which three different data sources of musical data are analyzed and clustered to find containment patterns and to discover musical dependencies also considering the lyrics of the songs.

3 Datasets and Preprocessing

In this section, we describe the two different types of dataset sources and the preprocessing applied to them.

As proxy of the actual musical scene, we exploit musical dataset having three different level of spatial granularity: world, national and regional. Datasets refer to *world*'s famous musicians, *Italian* musicians, and emerging youth bands in *Tuscany*, respectively.

In Table 1 we describe the details of these datasets. In particular, the `TUSCANY` dataset is referred to emerging artists participating in the "100 Band" contest promoted by "Tuscan Region" and "Controradio" in 2015 [2]. These datasets are built using the Spotify API [3] and are composed by all the songs present on the platform for the selected artists. For each artist, we collect songs titles, album titles, and the number of *followers*. Furthermore, each song is identified by its title, artist, and *popularity* score, that is based on more is played that song. In addition, each track is described by a set of musical features [7]: *acousticness,*

Table 1. Datasets statistics. Within brackets are reported the number of artists for which at least a single song lyric was available.

Dataset	#Artist	#Tracks	#Genres	#Lyrics
`WORLD`	833,197 (19,218)	5,525,222	1,380	79,204
`ITALY`	2,379 (710)	502,582	126	28,582
`TUSCANY`	513 (58)	24,147	28	91

danceability, duration, energy, instrumentalness, liveness, loudness, speechiness, tempo and *valence*. All the features range in [0, 1] with the exception of *duration, loudness* and *tempo*. In the preprocessing phase, we normalize these latter features in order to align all the feature scales.

Moreover, in order to deepen our analysis and consider only the mood transmitted by the players with their sing in terms of lexical content, we integrate these datasets with the lyrics of the songs. In Table 1 are also described the details of the lyrics' datasets. In particular, the `WORLD-lyric` dataset is collected from the Genius [4], the `ITALY-lyric` dataset is collected using SoundCloud API [5], and, finally, the `TUSCANY-lyric` is built extracting texts from the results of a survey. Using Google Form service [6] we gathered musical and personal data about artists who participated at the Tuscany 100 Band contest.

The first problem we decided to deal with is the grouping of musical genres. In fact, all three datasets show a large number genres (both minor and major). Using a list of popular music genres [8,10] we assign each song's minor genres to their major class. Then, we group the collected songs in 12 classes: country, blues, religious, hip hop, latin, electronic, folk, jazz, rock, R&B, pop, a cappella.

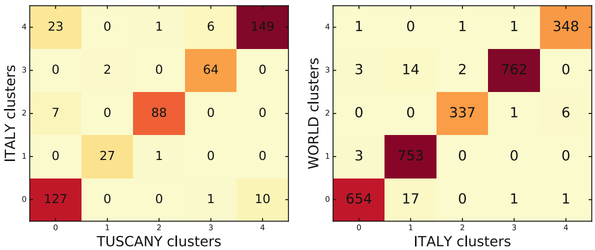
In order to process music lyrics with unsupervised methods we first clear the data. Therefore, to obtain normalized texts, we treat lyrics datasets using a rule-based cleaning procedure. Following this method, we obtain standardized music lyrics that can be treated by a general-purpose pipeline of Natural Language Processing (NLP). After that, music lyrics are lemmatized and Part-Of-Speech are tagged with the POS tagger TreeTagger [28,29]. We also reduce the noise selecting only nouns, verbs and adjectives. Doing this, for each text, we obtain only significant words from the sentiment points of view.

4 The Music Scene Fractal Structure

In order to identify the prototypical type of music produced by each artist in the datasets, we describe every performer through his medoid song, i.e., his most representative song identified minimizing the sum of the Euclidean distances between the Spotify features among all his discography. Once built such descriptions of each artist, we move on grouping them together on the basis of the music they produce. Since the available datasets allow us to observe the music phenomenon from three different hierarchical levels (regional, national and world-wide), we perform three different levels. The first one describing the music of regional artists, the second one describing the music of both regional and national artists and the last one describing the music of all the artists observed world-wide.

Through the analysis of the hierarchical clusterings, we aim at understanding if a fractal structure emerges among the type of music produced at different geographical levels. We accomplish this task by employing the *k-means* clustering algorithm [30] to the computed artist profiles [31]. As first step, in order to identify a reasonable value of k, i.e., the most appropriate number of clusters, we calculate the Sum of Squared Error (SSE) distribution for $k \in [2, 18]$. The SSE distributions for the clusterings computed on the three levels - not showed for space reasons - follow a common pattern that identifies optimal values of k in the range $[4, 6]$. After that we have identified such range, we extract the clusters for each value of k in it for the three datasets. Finally, from each cluster of each level, we calculate the medoid of the cluster, i.e., the set of features describing its most representative artist.

In order to understand if our datasets present a fractal structure, we study artists' migration among the clusters when moving from the regional to the world level. We repeat this activity for each k in the identified range. Figure 1 shows the clusters coverage - due to space reasons, only for $k = 5$ - when migrating from `ITALY` to `WORLD` clusters (right figure), and from `TUSCANY` to `ITALY` clusters (left figure). Indeed, these matrices have a strong diagonal prevalence. The same effect can be observed for all the clusters and for all k-values in the range [4, 6]: artists' blocks vary in high percentages from a down-geographical cluster to a top-geographical cluster.

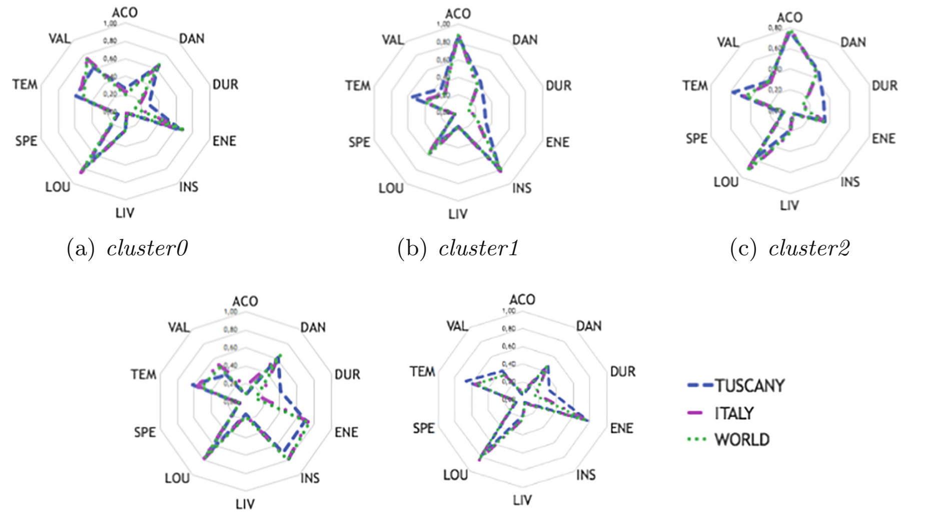


(a) k=5 (b) k=5

Fig. 1. Matrices clusters coverage when migrating from `ITALY` to `WORLD` clusters (left figure), and from `TUSCANY` to `ITALY` clusters (right figure).

We calculated such matrices by matching the pairs of clusters of two different levels with the highest level of coverage, i.e., by maximizing the purity using the cluster identifiers as a label. In fact, we can observe how, for instance, regional artists of a given cluster are re-classified into a cohesive block of the national level that, in turn is re-classified in a single block of the world level.

Since a fractal structure emerges with all the adopted values of k in the following we detail the clustering obtained only for $k = 5$. Moreover, music genres'



(a) *cluster0* (b) *cluster1* (c) *cluster2*

(d) *cluster3* (e) *cluster4*

Fig. 2. Artists Profiles: `TUSCANY`, `ITALY`, `WORLD` medoids. From left to right: *cluster0, cluster1, cluster2, cluster3, cluster4*

distribution varies within each of the three datasets, consequently affecting their clusterization. Thus, for each level, entire musical sub-genres fall into specific clusters.

In order to better analyze and understand the characteristics of these clusters, we compare the clusters for each dataset by taking advantage of the radar chart in Fig. 2 computed for $k = 5$ (we obtain comparable results for $k = 4$ and for $k = 6$). They describe the medoids of the five clusters identified for each dataset: radar charts underline, one more time, the presence of a fractal structure capturing very similar profiles across the observed hierarchy levels. This demonstrates that, for each different level of observation, musically homogeneous artists are well clustered with artists of the superior step. First of all, we notice how the *fractal* structure is perfectly highlighted also by the radar charts. The spikes of the three different datasets follow the same shape almost for each cluster. At a first glance of Fig. 2, we can observe that some features are more discriminant than others. In fact, features like *speechiness, liveness, loudness*, and *tempo* present similar values in each cluster. Despite this, the others features are determinants for cluster discrimination. Despite some little discrepancy among datasets, we can group clusters by their similarities. We perceive that *cluster0, cluster3* and *cluster4* are expressions of artists in direct opposition to the ones respectively in *cluster1* and *cluster2*:

– *cluster1* and *cluster2* represent melodic, bit danceable without a strong beat, and negative songs. Clusters diverge only for instrumental scores: *cluster1* present low values, on contrary *cluster2* show high values;

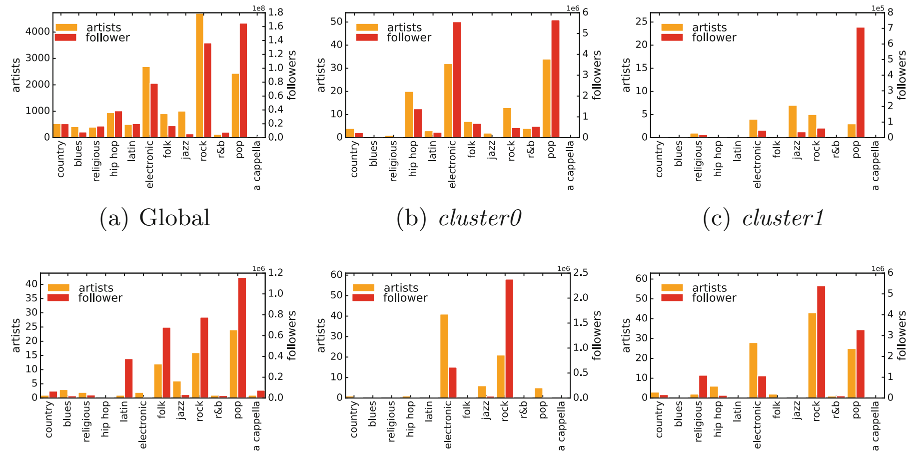
– *cluster0*, *cluster3* and *cluster4* represents non-melodic, strongly rhythmic and fairly dancing tracks. *cluster0* and *cluster4* show very low instrumental values, while *cluster3* show high scores. Furthermore, clusters differs for valence values: *cluster0* presents the highest values, on contrary *cluster4* presents the lowest values.

Since a fractal structure able to relate different geographical levels emerges also on the other dimensions, in the following sections we analyze the popularity, followers, genres, and sentiment of the clusters and we will detail only the results observed at the worldwide level.

5 Genres, Popularity and Followers

In this section, we analyze the collected information regarding artist popularity and followers as well as song genres. Starting from a dataset-wide discussion we detail how such dimensions can be used to characterize the identified clusters.

Figure 3(a) shows the overall distributions of both artists and followers with respect to the 12 genres identified after the cleaning stage. Indeed, the most represented genres are rock, electronic and pop, meta-classes. As was foreseeable, these genres also attract a considerable number of followers. However, is interesting noticing that, despite rock ranks a first among the most played genres, it presents a smaller number of followers than pop music.



(a) Global (b) *cluster0* (c) *cluster1*

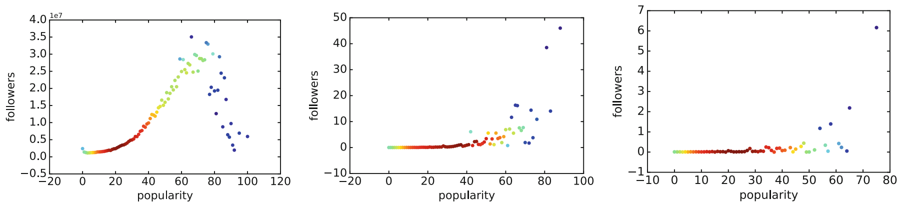
(d) *cluster2* (e) *cluster3* (f) *cluster4*

Fig. 3. Artists and Followers distribution among clusters.

Figure 4(a) shows the relation between artists' popularity and followers. Starting from such plot three subclasses of artists can be identified:

– *Low popularity.* A large number of artists present a low popularity, between 0 and 40, and a small number of followers: indeed, a large number of artists are followed by few people. Such artists could be emerging artists or are likely to play some kind of niche music.
– *Medium-high popularity.* Artists having medium popularity, between 40 and 70, are followed by a consistent number of users.
– *High popularity.* Artists having the very high popularity and that have a relatively low number of followers.

While the former two classes were somehow expected, the latter one breaks the common intuition that expects very popular artists to be the ones attracting more followers.



(a) Global (b) *cluster0* (c) *cluster1*

Fig. 4. Relations between popularity and followers.

Once observed the general behaviors of popularity, followers and genre components, we study how they relate with the clustering we obtained in Sect. 4. As a first step, in order to provide a semantic annotation of the identified clusters, we describe them exploiting their main genres as well as their profiles. As we can see from Fig. 3, clusters are strongly heterogeneous since they represent different music genres. Indeed, the clusters obtained are markedly different among each other and each cluster distinctly identifies a subset of genres with specific levels of popularity and number of followers.

Cluster0, Fig. 3(b), can be identified as the electronic/pop/hip hop cluster. It is represented by artists like Laura Pausini and Vanilla Sky. In this cluster fall few genres, however, musicians that belong to this profile are followed by the highest amount of users w.r.t. all clusters. Songs are very suitable for dancing, repetitive, cheerful and sung. In this cluster also fall all few rap artists but probably, they have little influence on medoids' values. However, speechness values are still the highest of the dataset.

Cluster1, Fig. 3(c), is the jazz/rock cluster. The majority of the artists belonging to this cluster, like Stefano Bollani and Doctor Dixie Jazz Band, have few followers. The beat pace and high acoustics make tracks unsuitable for dancing and influence the valence values that are lowest of all others clusters;

Cluster2, Fig. 3(d), is the pop/folk music cluster. These genres are the most represented by artists. In this cluster falls artists like Norah Jones and Lucio Battisti. Tracks are perceived as calm, unsuitable for dancing, primarily vocal and sad, thus leading to valences score that are fairly negative;

Cluster3, Fig. 3(e), is the electronic/rock cluster and it is represented by artists such Calvin Harris and Go!Zilla. The most representative genre is the electronic music, however, users that follow this genre are a few compared by the rock ones. This cluster represents non-acoustic, strongly danceable with a strong beat, like dance, house, and minimal music.

Cluster4, Fig. 3(f), is the pop/rock cluster and in this cluster falls artists like U2 and Green Day. The number of rock and pop followers is moderately high, while the amount of electronic music is quite low. Songs are strongly rhythmic, noisy and energetic but slightly danceable. Often, tracks are perceived as angry, so valences are tendentially negatives.

As final step we study for each cluster the relations between its artists' popularity and their number of followers. Due to space reasons, Fig. 4 shows the relation between artists' popularity and followers only for *cluster0* and for *cluster1*, but same results are showed for all five clusters. All the clusters are characterized by the same trend: most of their artists have medium-low popularity, with scores between 20 and 40, and are followed by a low amount of users. Moreover, as popularity grows also followers increase. However, musicians having both high popularity, over 70, and a high number of followers are few and almost uniformly spread across all the clusters. Moreover, it's interesting to observe that in each cluster there are very few artists with a very high popularity score, over 80, that are followed by a high number of followers: supposedly, these are famous international artists.

6 Sentiment Analysis

As a final analysis, we are interested in observing the correlation between the songs and the feelings they transmit, framing our analysis within our clusters.

We proceed by adopting a lexicon-based approach, exploiting ANEW [32] as seed-lexicon resource. ANEW provides a set of emotional ratings for 1,034 words in terms of *valence*, *arousal*, and *dominance*. In order to determinate artists' polarity, we select the valence values provided from both male and female subjects.

After calculating the weighted average of the words' valences v_{text} as polarity score, we grouped songs by artists, in order to obtain a polarity value for each of them. For each artist, we compute the weighted average among his tracks' polarity. In order to enhance the differences among the levels of the various scores we apply a logistic function [35] $f(x) = \frac{L}{1+e^{-k(x-x_0)}}$; where L is the curve maximum value (we set it equal to 1), $k = 10$ is the steepness of the curve, $x_0 = 0.5$ is the x-value of the sigmoid's midpoint, and $x = v_{text}$.

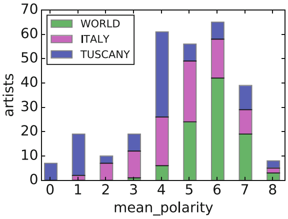


Fig. 5. Artists' polarity score distribution among polarity class.

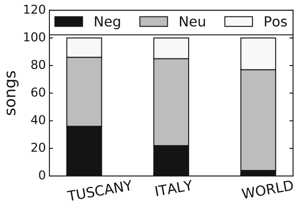


Fig. 6. Artists' polarity score distribution among datasets.

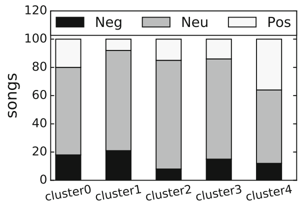


Fig. 7. Artists' polarity score distribution among clusters.

We then apply the aforementioned procedure to each lyrics' dataset. Then we grouping emotionally-tagged tracks for each artist and computing the weighted average among the i-artist's discography. In order to analyze text belonging to `ITALY-lyric` and `TUSCANY-lyric` we translated the ANEW lexicon in Italian by using the Python library Goslate [9].

As result, we obtain a comprehensive list of artists each one of them having a polarity scores in the range [0, 10]. Studying the polarity distributions we note that the most artists have a polarity score in the range [5, 7], as showed in Fig. 5. Anyway, we need to keep in mind that a neutral score could indicate that *(a)* the artist's tracks transmit no strong emotions or, *(b)* they present conflicting emotions. In order to give a better comprehension of the results, we split artists based on the polarity score into three class: *(a) positive*, scores higher than 6; *(b) negative*, scores lower than 4, and *neutral*, scores between 4 and 6. Figure 6 shows the distribution of the artists' polarity among the three datasets using ANEW.

Finally, we analyze how the clusters are affected by the polarity scores. Indeed, we apply the same procedure described above for each of the five clusters. The other clusters are described by a Gaussian curve similar to the one expressed by the complete dataset. In fact, all such distributions are multimodal and define three peaks. The remaining artists reach more extreme polarity scores, with a long tail to negative values.

Going further into the analysis, we split the artists based on their scores into the three polarity class identified above. Figure 7 shows the distribution of the artists' polarity among the three clusters using ANEW. As we can observe, the most represented category still remains the neutral one, while the less represented is the negative one.

7 Conclusion

In this work, we have proposed a data-driven investigation of the music scene. We have fulfilled this task by relying on a composite dataset built upon heterogeneous on-line resources. We compared song technical features, lyrics and artist popularity across three hierarchical geographic layers (world, national and regional). Our analysis identifies the existence of very stable clustering structure able to describe cross-genre music profiles. We highlighted how such clusters describe a *fractal* structure: apparently, disregarding the geographical granularity observed, all the artists observed can be profiled and categorized in a reduced and well defined set of clusters. Moreover, we analyzed artists' popularity and fan base observing how their distributions describe a similar trend in all the identified clusters. Finally, looking to the artists' song lyrics we were able to observe the emotional valence of the identified meta-profiles.

As future works we plan to consolidate (and extend with data coming from other sources) the cross-domain dataset we collected and to build upon it – exploiting the results of the present study – a recommender system that enables its users to discover novel artist on the base of their tastes, mood and locations.

Acknowledgment. This work is partially supported by the European Project *SoBigData: Social Mining & Big Data Ecosystem*, http://www.sobigdata.eu. This work is partially supported by the European Communitys H2020 Program under the funding scheme "INFRAIA-1-2014-2015: Research Infrastructures" grant agreement, http://www.sobigdata.eu, GS501100001809, 654024 "SoBigData: Social Mining & Big Data Ecosystem".

References

1. Spotify. https://www.spotify.com/. Accessed 28 July 2017
2. Toscana100band contest. http://toscana100band.it/. Accessed 28 July 2017
3. Spotify Web API. https://developer.spotify.com/web-api/. Accessed 28 July 2017
4. Genius. https://genius.com/. Accessed 28 July 2017
5. SoundCloud API. https://developers.soundcloud.com/docs/api/guide. Accessed 28 July 2017
6. Google Form service. https://www.google.com/forms/about/. Accessed 28 July 2017
7. Spotify Audio Features Object. https://developer.spotify.com/web-api/get-several-audio-features/. Accessed 28 July 2017
8. AllMusic. http://www.allmusic.com/genres. Accessed 28 July 2017
9. Goslate. http://pythonhosted.org/goslate/. Accessed 28 July 2017
10. List of popular music genres, (n.d.). Wikipedia. http://en.wikipedia.org/wiki/Psychology. Accessed 28 July 2017
11. Clarke, F., Ekeland, I.: Nonlinear oscillations and boundary-value problems for Hamiltonian systems. Arch. Rat. Mech. Anal. **78**, 315–333 (1982)
12. Clarke, F., Ekeland, I.: Solutions périodiques, du période donnée, des équations hamiltoniennes. Note CRAS Paris **287**, 1013–1015 (1978)
13. Michalek, R., Tarantello, G.: Subharmonic solutions with prescribed minimal period for nonautonomous Hamiltonian systems. J. Differ. Equ. **72**, 28–55 (1988)
14. Tarantello, G.: Subharmonic solutions for Hamiltonian systems via a $\mathbb{Z}_p$ pseudoindex theory. Annali di Matematica Pura (to appear)
15. Rabinowitz, P.: On subharmonic solutions of a Hamiltonian system. Commun. Pure Appl. Math. **33**, 609–633 (1980)
16. Guidotti, R., Rossetti, G., Pedreschi, D.: Audio ergo sum. In: Milazzo, P., Varró, D., Wimmer, M. (eds.) STAF 2016. LNCS, vol. 9946, pp. 51–66. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-50230-4_5

17. Rawlings, D., Ciancarelli, V.: Music preference and the five-factor model of the NEO personality inventory. Psychol. Music **25**(2), 120–132 (1997)
18. Rentfrow, P.J., Gosling, S.D.: The do re mi's of everyday life: the structure and personality correlates of music preferences. J. Pers. Soc. Psychol. **84**(6), 1236 (2003)
19. Bu, J., Tan, S., Chen, C., Wang, C., Wu, H., Zhang, L., He, X.: Music recommendation by unified hypergraph: combining social media information and music content. In: Proceedings of the 18th ACM International Conference on Multimedia, pp. 391–400. ACM (2010)
20. Bischoff, K.: We love rock'n'roll: analyzing and predicting friendship links in last.fm. In: Proceedings of the 4th Annual ACM Web Science Conference, pp. 47–56. ACM (2012)
21. Pennacchioli, D., Rossetti, G., Pappalardo, L., Pedreschi, D., Giannotti, F., Coscia, M.: The three dimensions of social prominence. In: Jatowt, A., Lim, E.-P., Ding, Y., Miura, A., Tezuka, T., Dias, G., Tanaka, K., Flanagin, A., Dai, B.T. (eds.) SocInfo 2013. LNCS, vol. 8238, pp. 319–332. Springer, Cham (2013). https://doi.org/10.1007/978-3-319-03260-3_28
22. Putzke, J., Fischbach, K., Schoder, D., Gloor, P.A.: Cross-cultural gender differences in the adoption and usage of social media platforms - an exploratory study of Last.FM. Comput. Netw. **75**, 519–530 (2014). http://dx.doi.org/10.1016/j.comnet.2014.08.027
23. Park, M., Weber, I., Naaman, M., Vieweg, S.: Understanding musical diversity via online social media. In: Ninth International AAAI Conference on Web and Social Media (2015)
24. Zheleva, E., Guiver, J., Mendes Rodrigues, E., Milić-Frayling, N.: Statistical models of music-listening sessions in social media. In: Proceedings of the 19th International Conference on World wide web, pp. 1019–1028. ACM (2010)
25. Li, T., Ogihara, M., Peng, W., Shao, B., Zhu, S.: Music clustering with features from different information sources. IEEE Trans. Multimed. **11**(3), 477–485 (2009)
26. Peng, W., Li, T., Ogihara, M.: Music clustering with constraints. In: ISMIR, pp. 27–32 (2007)
27. González-Pardo, A., Granados, A., Camacho, D., de Borja Rodríguez, F.: Influence of music representation on compression-based clustering. In: 2010 IEEE Congress on Evolutionary Computation (CEC), pp. 1–8. IEEE (2010)
28. Schmid, H.: Improvements in part-of-speech tagging with an application to German. In: Proceedings of the ACL SIGDAT-workshop. Citeseer (1995)
29. Schmid, H.: Part-of-speech tagging with neural networks. In: Proceedings of the 15th Conference on Computational Linguistics, vol. 1, pp. 172–176. Association for Computational Linguistics (1994)
30. Tan, P.-N., Steinbach, M., Kumar, V., et al.: Introduction to Data Mining, vol. 1. Pearson Addison Wesley, Boston (2006)
31. Pollacci, L., Guidotti, R., Rossetti, G.: Are we playing like Music-Stars? Placing emerging artists on the Italian music scene (2016)
32. Bradley, M.M., Lang, P.J.: Affective norms for English words (ANEW): instruction manual and affective ratings. Citeseer, Technical report (1999)
33. Lang, P.J.: Behavioral treatment and bio-behavioral assessment: computer applications (1980)
34. Dodds, P.S., Danforth, C.M.: Measuring the happiness of large-scale written expression: songs, blogs, and presidents. J. Happiness Stud. **11**(4), 441–456 (2010)
35. Verhulst, P.F.: Recherches mathématiques sur la loi d'accroissement de la population. Nouveaux mémoires de l'académie royale des sciences et belles-lettres de Bruxelles **18**, 14–54 (1845)

Usable and Accessible Tourism Websites for Children: A Case Study of a Naturalistic Oasis

Antonina Dattolo[1(✉)], Chiara De March[2], and Flaminia L. Luccio[3]

[1] SASWEB Lab, DMIF, University of Udine, Gorizia, Italy
antonina.dattolo@uniud.it
[2] Università Ca' Foscari Venezia, Venezia, Italy
855411@stud.unive.it
[3] DAIS, Università Ca' Foscari Venezia, Venezia, Italy
luccio@unive.it

Abstract. The children's use of technology is increasing year by year and recent statistics show that children receive their first smartphone and autonomously navigate the Web when they are 10.3 years old (on average). Nevertheless, while the number of tourism packages for families with children grows, the number of websites that can be directly visited by children is small; more relevant, is the number of tourism applications developed for children. This paper discusses the design of usable and accessible websites for children, and proposes, as a case study, our experience on a dedicated section of the tourism website of a Naturalistic Oasis located close to Treviso, in Italy.

1 Introduction

A recent study, conducted in 2016 by the Influence Central [12], revels interesting statistics about the use of technologies by children in North America: children start receiving their first smartphone at 10.3 years of age, and they use it mainly for texting people and parents even in the same environment, e.g., same room, same building, etc. These tools have become a sort of their own physical extension. During trips, tablets are in 55% of the cases the preferred entertainment tools, then phones in 45% of cases.

Regarding the access to Internet, the parents tend to be more permissive than in the past: only 76% of them impose strict limits on the independent use of the connection, compared to 85% in 2012.

An interesting survey, conducted online in 2016 by the Erikson Institute, shows that in the United States even children below 6 years of age have a strong contact with technologies: parents reported that 85% of their children had a contact with television, tablets, smartphones, and computers on a daily base for up to two hours [11].

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 195–205, 2018.
https://doi.org/10.1007/978-3-319-76111-4_20

On the basis of these facts, two new challenges open up for the tourism:

- for *tourism companies*, as creating new digital cultural and tourism material addressed also for young users might be a winning strategy when the tourism destinations are devoted to families or schools. Tourism websites should provide recreational and educational material personalised to children, and are meant both to enhance the cultural and naturalistic heritage, and also to support practical and informative material for the tourism experience. Children will not be clients of the tourism experience, they will not be part of an economic exchange, but will be direct consumers. An involved and happy child will improve the tourism experience of the whole family and parents will become satisfied clients.
  Different studies show how children may influence the choice of a tourism activity [17]. A child that enjoyed the website may ask the parents to visit the described place, and in turn also a parent may find in this nice website a sign of attention to the recreational and educational needs of a child, and thus a better quality of the proposed tourism experience.
- while dealing with *educational tourism*, for which the target consumers are students at school. Proposing interesting and well structured material online may be interesting to improve the didactical experience in class before, and after the tourism tour, and may also increment the interest of teachers, thus improving the competitiveness of the related tourism location.

Nevertheless, while the number of tourism packages for families with children and accessible tourism locations for children grows (see for example Pantou.org), the number of usable and accessible (U&A) tourism websites for children is still small; more relevant is the number of tourism apps developed for children, such as [1,3], which are designed to entertain children in an intelligent and cultural way, encouraging them to ask questions and to think.

In the meantime, the Advocate for Children and Young People strives to maintain conformance to W3C's Web Content Accessibility Guidelines (WCAG); the UNTWO (World Tourism Organization) and the ENAT (European Network for Accessible Tourism) emphasise the relevant role of an inclusive Web design for tourism.

In this paper, we would like present our experience and work in designing and realising a section (for children) of the tourism website of a naturalistic oasis, called *Oasi Cervara*, close to the city of Treviso, in the North of Italy.

The paper is organised as follows: Sect. 2 introduces the concepts of usability and accessibility, explores the state of art and related work regarding the realisation of a tourism website for children, and proposes general guidelines for its development. Section 3 presents our case study, focused on a U&A tourism, children-friendly website. Section 4 describes our experimental results, and finally Sect. 5 ends the paper with a look to future work.

2 Website Usability and Accessibility

Usability and accessibility are closely related and they overlap significantly; in the design of a website is effective to address them together. Following the definitions provided in [20]: usability is about designing products to be effective, efficient, and satisfying, while accessibility addresses discriminatory aspects, including people with age-related impairments. In this paper, we consider only this aspect of the accessibility. For general accessibility, guidelines we refer the reader to [16].

2.1 State of Art

Making U&A for children a website that already owns these features for adults does not just consist in adding some colours, puppets and in eliminating or simplifying actions that adults consider difficult [18]. The same considerations apply for children of different ages, as older children show more experience than younger ones. This is the result of an interesting study that was conducted in 2010 by Jacob Nielsen using two rounds of testing: the first on 27 sites with 55 children, aged 6–11 in 2001; the second on 29 sites with 35 children, aged 3–12 years in 2010 [15].

Both studies, conducted at 9 years of distance, provide similar results with the only difference that the early exposure to computers has anticipated the abilities that children showed around 6–8 years to 3–5 years.

In [15], the author addresses different key features that have to be taken into consideration while developing sites for users aged 3–12 years. Note that, other guidelines have also been proposed for teenagers [14]. For users in the age range 3–12, the first thing is to divide them in three groups young (3–5), mid-range (6–8), and older (9–12), as following specific guidelines for each group improves the user experience. Then, it is important to keep in mind that adults and children have different needs/perceptions.

From a usability study conducted in [13], it emerges the importance of providing a simple but funny name to the site to attract the child's attention. Also the related link should be easily memorised for future visits. The initial impression is also very important as it affects the whole user experience [18]. Moreover, during navigation, the user wants to have a full control of the site in order to be gratified, thus the navigational structure should be similar on each page of the site and should be easy to follow [19]. The search for items should be simple and direct [13]. Moreover, children want instant feedback for every action they make [21]. Having technical limitations provided, e.g., by the need to install plugins or update software, or by the access to old desktops/laptops, etc. may negatively affect the user experience. The same holds for mistakes that should be easily recovered from, or for promotional content that typically attracts the child but at the same time distances him from the site. To be enjoyable for a child, the site should offer entertainment such as games or interactive material. Personalisation of colour choices, graphical devices and animations is also considered important [13]. The text should be easy and succinct, at an appropriate readability level for the age range, colors should be bright and with a good contrast, the background should be colourful and not white [13]. It should still be clear which elements of the site are interactive and which are not [4,8,21].

Table 1. Topics of interest of the children's websites.

Website/Feature	Animals	Badges	Coloring	Didactic activities	Games	Quizzes	Reading	Videos
PBS kids				√	√	√	√	√
Sesame street			√	√	√			√
Starfall	√			√	√		√	
CoolMath4kids		√		√	√	√		
MakeMeGenius	√			√	√	√	√	√
National geographic kids	√	√		√	√	√		√
Fun brain		√		√	√		√	√
Nick Jr.			√	√	√		√	√
Disney Jr.		√	√	√	√			√
Highlights for kids			√	√	√	√	√	√
Total	30%	40%	40%	100%	100%	50%	60%	80%

2.2 Related Work and Guidelines

The number of tourism websites for children is still small. Very recently, the *Independent* published a list of the 50 best travel websites [2]: it appears evident that all of them directly address adults and only indirectly children. A recent list of the 20 of the best free and fun educational websites for children [10] provides an idea of the widest topics of interest of the current portals dedicated to children. In Table 1 we list the 10 websites that apply U&A guidelines and provide us an idea of the topics, considered appealing for children. We can note that the topics of interest may be considered general modalities for capturing the attention of the children; furthermore, these websites share general guidelines regarding the graphical layout, the navigation structure, the used language and the user model. Keeping in mind the common features of these websites, the previous observations that we mainly collected from [15], but also from [4–8,13,18,19,21], we provide a list of general guidelines for a U&A tourism website:

- The **design** and the **structure** should be simple, clear and *predictable*; the content focused on the objectives of the website. A *fun name/logo* should grab the user attention; the URL should be a predictable one. The information should be redundant, and a graphical visual interface should contain a copious use of pictures (drawings, photographs, symbolic images), easy to understand. The text should go with pictures. It should be *clear, simple, and short*; should be in a big font (14 characters points for younger and 12 points for older children), in plain Sans-serif style (e.g., Verdana), with bright colours and a good contrast. Headings and titles should be used.

The interface should be *responsive* and *perceivable*, for enabling the use of screen readers. To avoid: background sounds, moving text, blinking images and horizontal scrolling.
- **Navigation** should be consistent, similar in every page/section, and propose a simple and *logical structure* of navigation, also using navigation *buttons* at the top and the bottom of the page, and back/forward buttons. The website should be equipped with *search engines*.
- The website should **engage the users**, providing them entertaining, quick to judge (standard), clear, consistent, controllable, instantly *gratifying*, with feedback contents; should introduce the *gamification*; should allow *customisation* and make adaptive the interaction with users, considering their interaction history, their preferences, requests, needs, and their diversified goals; should propose the tasks decomposed in simple subtasks.
- The **language** should be simple and precise, and apply the adequate linguistic registers for the *target age*, supplied as necessary with *animations and sounds*; acronyms and abbreviations, non-literal texts, and jargon should be avoided.

Furthermore, for what concerns tourism children websites we obviously point out that all the previous suggestions apply as the site should be enjoyable for children. In this context we want to emphasize the key points that we have tested also in our case study that will be presented in Sect. 3:

1. Providing interesting touristic material targeted for children, i.e., with a short interesting text whose reading creates curiosity and expectations;
2. addition of interactive games containing information related to the target destination.

That is, the site should follow standard children usability rules, but at the same time should create curiosity to visit the tourism destination.

3 Our Case Study: A Naturalistic Oasis

In this section we describe our case study. For a naturalistic oasis called *Oasi Cervara*, we created a section of the website, called *Oasi Kids*[1], designed for the direct navigation by children [9].

The Oasis is situated in the Veneto Region in the North of Italy and has a special environment composed of a swamp ambient and a peat bog system, and provides habitat for different spices of animals; for this reason is considered by the Veneto Region as Community Interest Site and Special Protection Zone for the fauna of the Sile river.

The Oasi Cervara S.r.l. manages the park trying to combine the protection and conservation of the ambient together with visitors' enjoyment. This park is mainly visited by families (in 79% of the cases), schools (in 19% of the cases), photographers (in 2% of the cases), or elderly people. The Oasis has a certificate

[1] Section temporarily offline from the official site http://oasicervara.it/ for technical reasons, but part of it is online at this link: http://oasikids.altervista.org/oasi-kids/.

of Excellence on TripAdvisor; however the few negative scores indicate that there is a gap between the expectations (e.g., seeing a lot of animals), and what it is really offered (which obviously also depends on the season, the weather, etc.). Thus, improving the quality of the offered information on the website is a challenging but important point, and for this reason, we proposed and developed the *Oasi Kids* section, that has the aim of:

1. being inclusive, usable and accessible for children in the range 8–11 years. The site contains some descriptive parts not suitable for younger users;
2. enriching the online communication of the information about the Oasis;
3. improving the quality perceived by families and teachers of the offer proposed by the Oasis.

The site has been developed using Wordpress, following the guidelines presented in previous Sect. 2.2, in order to be inclusive for the chosen target. In particular, regarding accessibility, the site is:

– *Responsive*: may adapt to mobile phones, to different screens and browsers;
– *Perceivable*: following the W3C guidelines the language has been specified, and there is alternative text for screen readers;
– *Distinguishable*: there is a good contrast between the colours of pages and of the text, both inside the menu (black over white), and also inside the standard pages (white over grey/black);
– *Compatible*: the choice of a standard font Arial and of the font-family sans-serif, guarantees the maximal compatibility with all the browsers that may access the site.

Regarding usability:

– *Navigation*: is simple, the user is directed to a specific page via a menu and the related sub-menu or by clicking big buttons with images and text;
– *Structure*: very simple and functional;
– *Layout and graphics*: simple, targeted to young users;
– *Home Page*: adaptable by proposing different topics to different users;
– *Text*: the length is medium/short, some words are in a different color to emphasise a concept and to create curiosity;
– *Syntax* of the content is simple and clear.

Regarding the layout and the navigation, clickable areas, where placed with an image and text. E.g., in Fig. 1 there are round clickable areas in a shape of a flower to be more enjoyable by young users. There is also a button "Torna al Menu" (Go back to the menu) to simplify navigation. The depth of the navigation is at most 3 clicks. In Fig. 1 we may notice that the text is medium/short, and some words are in a different color recalling the concept, e.g., erba (grass) is in green, etc.

Regarding the content, the arguments have been chosen so to create curiosity, interest to visit the park (e.g., messages such as "did you know that...") , and at the same time contain precise but simple scientific information (such as the family name of an animal, the physical characteristics, etc.). To simplify the educational experience each animal card contains a pdf version that can be downloaded for pre and post-processing. Moreover, the site proposes educational messages, as shown, e.g., in Fig. 1 with the message "If you want to see the animals you should be quiet and careful...".

The used language is informal, simple and clear, the description of the animals is in first person. In some parts the text has been written using nursery rhymes, to make the content more enjoyable, musical and memorisable, and also to increase the curiosity of the child for the real visit to the park.

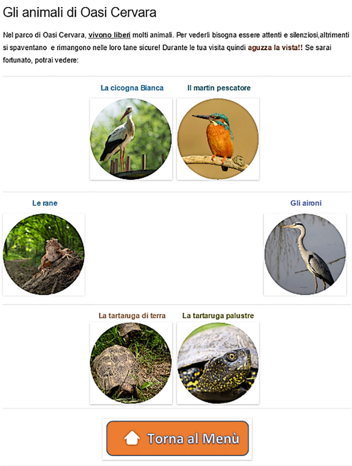


Fig. 1. Part of a page with animals. (Color figure online)

Fig. 2. A memory game with animals.

Finally, to improve the navigational and educational experience of the children, the Oasi Kids site proposes two games: a memory game with animal pictures (see Fig. 2), and some cards with questions related to the different animals described in the site. To avoid that young users get impatient during the game loading, they are entertained by an animation of an owl that open and closes its eyes.

4 Experimental Results

In this section we describe the *methodology* we have used and we then illustrate our experimental results on the Oasi Kids site. We have developed the website following different steps:

– **Literature survey:** In the first phase, we have collected all the information about usable and accessible websites for young users, and we have defined some features the site should include (e.g., many pictures, games, etc.).

- **Interviews:** In the second phase, in order to decide how to structure the site and which contents to include, we have conducted two different interviews inside a primary school. The first one with a group of 42 parents of children of the school, the second one with 30 teachers of the school, to check their possible interest for having a website dedicated to children on a tourism educational destination.

  Interviews to parents. The three questions were:
  1. The use of technology by children: 90.5% of the children use by their own technology for games/school research.
  2. The search of information for a tourism destination, and the involvement of a child in the search: 100% of the parents search tourism information on the Web, 55.33% completely involve their children in the decision, 36.6% involve them partially.
  3. The importance of the development of the Oasi Kids website: The average score was 6 out of 7.

  The parents were also asked about the material they would expect to find in the site, and the major requests were videos and photos (in 84.2% of the cases), information about the park (in 76.3%) and about the flora and fauna (in 71.7%). Finally, in 81.6% of the cases the parents claimed that while choosing between two tourism attractions having the same price, they would pick the one with a website section dedicated to children.

  Interviews to teachers. The four questions were:
  1. The use of technology by children outside school and at school: Children use technology outside school for games, school research, videos, communication with parents; at school for guided searches on the Web, didactic activities inside educational websites, watching videos dedicated to linguistic and musical activities.
  2. The set of activities proposed to children before a tourism didactic experience: Typically teachers try to prepare children to the experience by anticipating them the content of the visit with didactic presentations (37% of the cases), laboratories (25%), etc., and the way they should behave. This, in their opinion, increases the quality of the educational experience.
  3. The best pre-visit information tool: Technological tools with videos (24 out of 30 teachers), information material (27 out of 30), games (20 out of 30), etc. So all the teachers claimed they would use a site with such a content to prepare children to the visit, and 2 teachers out of 30 stated they would also use it as a post-processing tool after the visit, to let the students internalise and deepen the contents.
  4. The importance of the development of the Oasi Kids website: 93% of the teachers claimed it would improve the quality of the didactic experience.

  To conclude, for both interviews emerges the great importance of developing a dedicated website containing videos, photos, didactic material, and games.

- **Website development:** The third phase was the development of the website applying the usability and accessibility guidelines discussed in Sect. 2.2, and the hints and suggestions collected in the literary survey and interview phases.

– **Assessment questionnaire:** In order to evaluate the perceived impression of the Web site, in the last phase we collected through an assessment questionnaire the opinions from a set of 50 parents and separately from a set of 50 children of an age between 8 and 11 years.
The questionnaires for the parents had scores that ranged from 1 to 7, for children we have used a different scale composed of 4 different emoticons (representing the perception of: very good, good, so and so, do not like). The main results of the questionnaire proposed to the parents are summarised in Table 2, more details may be found in [9]. Regarding the questionnaire proposed to the children, the results are summarised in Table 3: The children liked the site in general, liked colours and images, the descriptions and they all would be exited to visit the Oasis.

Table 2. Results of the assessment questionnaire proposed to 50 parents.

Questions	Average grade
Do you like the Oasi Kids site?	5.6/7
Do you like the site content?	5.62/7
Do you like the site information?	5.54/7
Do you like the site colours and characters?	5.4/7
Was the quality of the tourism destination increased by the website?	4.9/7

Table 3. Results of the assessment questionnaire proposed to 50 children.

Questions				
Do you like the Oasi Kids site?	20	23	7	-
Do you like colors and images?	23	22	3	2
Was it easy to find the material you searched for?	14	26	10	-
Were the topic descriptions interesting?	22	20	6	2
Do you feel like visiting the Oasis now?	20	25	5	-

To conclude, the parents and the children have in general appreciated the website Oasi Kids and, although some time is needed to have a real feedback, in our opinion this could be a good starting point for the increment of visits to the Oasi Cervara park.

5 Conclusion and Future Challenges

In this paper, we analysed how tourism websites should be developed to be usable and accessible by children. We have shown how a child that enjoys the navigation on a site may influence the choice of a particular tourism destination by the whole family or by a school for didactic purposes. We have illustrated an interesting U&A children tourism website that we have developed following the proposed usability and accessibility guidelines, and we have discussed in details the result of an assessment questionnaire proposed to parents and to children, that shows how the site was well appreciated by both groups.

As a future work we are developing another tourism website on which we are planning to do a more extensive analysis by considering features such as users (children) experience, improved experience, increased number of visits or longer sessions. We also plan to propose a set of templates and widgets conceived specifically for tourism websites.

References

1. I leoni del tempo. Tourism app for children (2017). https://goo.gl/HBsfSb
2. Ahluwalia, R.: The 50 best travel websites, Indipendent (2017). https://goo.gl/cr3XK7
3. Art Stories: Journey Journal (2017). http://www.artstories.it
4. Bernhardt, G.: Designing Usable Sites for Children and Teens (2006). https://msu.edu/user/webbsuza/atw/graceb/
5. Dattolo, A., Luccio, F.L.: A review of websites and mobile applications for people with autism spectrum disorders: towards shared guidelines. In: Gaggi, O., Manzoni, P., Palazzi, C., Bujari, A., Marquez-Barja, J.M. (eds.) GOODTECHS 2016. LNICST, vol. 195, pp. 264–273. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61949-1_28
6. Dattolo, A., Luccio, F.L.: Accessible and usable websites and mobile applications for people with autism spectrum disorders: a comparative study. EAI Endorsed Trans. Ambient Syst. **4**(13) (2017)
7. Dattolo, A., Luccio, F.L., Pirone, E.: Web accessibility recommendations for the design of tourism websites for people with autism spectrum disorders. Int. J. Adv. Life Sci. **8**(3–4), 297–308 (2016)
8. De Cock, R., Hautekiet, E.: Children's news online: content analysis and usability study results (2011). https://goo.gl/FLSb8Z
9. De March, C.: La personalizzazione dei contenuti Web per migliorare l'attrattività delle imprese turistiche e culturali: il caso di studio di Oasi Cervara. Master's thesis, Università Ca' Foscari, Venezia (2016). https://goo.gl/5dAgzH
10. Ducan, A.: Educational websites for kids that are free, the spruce (2017). https://www.thespruce.com/best-free-educational-websites-for-kids-3129084
11. Erikson Institute: Technology and Young Children in the Digital Age (2016). https://goo.gl/hZSQCw
12. Influence Central: Kids & Tech: The Evolution of Today's Digital Natives (2016). https://goo.gl/6K8NgW
13. Large, A., Beheshti, J., Rahman, T.: Design criteria for children's web portals: the users speak out. J. Am. Soc. Inform. Sci. Technol. **53**(2), 79–94 (2002)
14. Loranger, L., Nielsen, J.: Teenage usability: designing teen-targeted websites (2013). https://goo.gl/81zY58
15. Nielsen, J.: Children's websites: usability issues in designing for young people (2010). https://goo.gl/D5q35b

16. Seeman, L., Cooper, M.: Techniques for the the cognitive and learning disabilities accessibility task force (COGA), W3C (2016). https://goo.gl/bdKhQi
17. Srnec, T., Lončarić, D., Perišić Prodan, M.: Family vacation decision making process: evidence from Croatia. In: Proceedings of the Congress on Tourism and Hospitality Industry, Opatija, Croatia - Hrvatska, pp. 432–445, 28–29 April 2016
18. UxKids: Comparing Usability for Kids and Adults (Part 1) (2013). http://uxkids.com/blog/comparing-usability-for-kids-and-adults-part-1/
19. Vučković, K., Librenjak, S., Dovedan, Z.: Internet for the youngest: computer usage and web design for the early school age children. Int. J. Emerg. Technol. Learn. **5**(S12), 17–23 (2010)
20. WAI: Accessibility, usability, and inclusion: Related aspects of a web for all (2016). https://www.w3.org/WAI/intro/usable
21. White, B.: Designing For Kids Is Not Childs Play (2016). https://goo.gl/nYx9bK

Parkinson's Disease Detection from Speech Using Convolutional Neural Networks

Evaldas Vaiciukynas[1,2](✉), Adas Gelzinis[1], Antanas Verikas[1,3], and Marija Bacauskiene[1]

[1] Department of Electrical Power Systems, Kaunas University of Technology, Studentu 50, 51368 Kaunas, Lithuania
{evaldas.vaiciukynas,adas.gelzinis,marija.bacauskiene}@ktu.lt
[2] Department of Information Systems, Kaunas University of Technology, Studentu 50, 51368 Kaunas, Lithuania
[3] Centre for Applied Intelligent Systems Research, Halmstad University, Kristian IV:s väg 3, PO Box 823, 30118 Halmstad, Sweden
antanas.verikas@hh.se

Abstract. Application of deep learning tends to outperform hand-crafted features in many domains. This study uses convolutional neural networks to explore effectiveness of various segments of a speech signal, – text-dependent pronunciation of a short sentence, – in Parkinson's disease detection task. Besides the common Mel-frequency spectrogram and its first and second derivatives, inclusion of various other input feature maps is also considered. Image interpolation is investigated as a solution to obtain a spectrogram of fixed length. The equal error rate (EER) for sentence segments varied from 20.3% to 29.5%. Fusion of decisions from sentence segments achieved EER of 14.1%, whereas the best result when using the full sentence exhibited EER of 16.8%. Therefore, splitting speech into segments could be recommended for Parkinson's disease detection.

Keywords: Parkinson's disease · Audio signal processing
Convolutional neural network · Information fusion

1 Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disease after Alzheimer's [1] and it is expected that the prevalence of PD is going to increase due to population ageing. Medical intervention could be considered to slow down the progression of PD if it is detected early, resulting in increased life span and life quality for PD patients. Acoustic analysis of voice or speech signal is considered as an important non-invasive tool in screening for PD. Related work is summarized in [2], where splitting of speech into voiced and unvoiced parts is recommended.

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 206–215, 2018.
https://doi.org/10.1007/978-3-319-76111-4_21

Recent advances in deep learning helped reaching the state-of-the-art performance in various domains – not only for images, but for audio data as well. For example, the combination of deep neural networks with hidden Markov models (DNN-HMM) in [3] outperformed traditional the Gaussian mixture model-based (GMM-HMM) solution. Convolutional neural networks (CNN) were successfully applied to automatic speech recognition [4,5], speech activity detection [6,7] or acoustic scene classification [8]. When applying CNN to audio data it is common to characterize an underlying signal using spectrograms, obtained by calculating the Mel-frequency spectral coefficients (MFSCs) from extracted fragments of a signal, as in [5,6,9,10]. Besides spectrograms, their first and second temporal derivatives (delta and double delta) can also be considered as additional input feature maps [4,7,8,11].

This study explores speech recordings of a four-words sentence in Lithuanian language processed by the spectrogram-based CNN model for the task of PD detection. Splitting a sentence into various segments, corresponding to separate words and combinations of words or syllables, is considered. Besides MFSCs and their first and second derivatives, usage of various other input feature maps is also proposed. Due to different lengths of speech segments, corresponding to the same part of the sentence, a solution to obtain a fixed length spectrogram by image interpolation is compared to the commonly used sampling of a fixed size window at random locations. Decision-level fusion is applied to improve PD detection.

2 Data

Pronunciation of a phonetically balanced sentence in a native Lithuanian language "turėjo senelė žilą oželį" (which translates into "granny had a little greyish goat") was recorded in a sound-proof booth. Recordings were done using an acoustic cardioid (AKG Perception 220, frequency range 20–20000 Hz) microphone. Microphone was located at ∼10 cm distance from the mouth. The audio

Table 1. Sentence segments, containing separate words (# 1–4), transitions between words when split on syllables (# 5–6), pairs of words (# 7–8), or full sentence (# 9).

#	Sentence fragments				Notation
1	tu rė jo				TUREJO
2		se ne lė			SENELE
3			ži lą		ZILA
4				o že lį	OZELI
5		**lė**	ži lą	**o**	LE_O
6	**rė** jo	**se**			RE_SE
7	**tu** rė jo	se ne **lė**			TU_LE
8			**ži** lą	o že **lį**	ZI_LI
9	tu rė jo	se ne lė	ži lą	o že lį	SENTENCE

format was mono PCM wav (16 bits at 44.1 kHz sampling rate). A mixed gender database collected contains 268 subjects (194 healthy controls and 74 PD cases) ranging from 22 to 85 years in age.

Each speech recording was manually annotated and split into sentence segments, containing: separate words, transitions between words and pairs of consecutive words. A full sentence without any splitting (*SENTENCE* segment) was also considered. Details on sentence segments used are in Table 1.

3 Methodology

3.1 Input Feature Maps

An important step in acoustic analysis is characterization of an audio signal by various features. Mel-frequency spectral coefficients (MFSC) is a commonly applied transformation to audio signal spectrum, resulting in the Mel-warped spectrogram. The MFSC spectrogram contains values of amplitude for each frequency coefficient (on the vertical axis) and a time moment (on the horizontal axis). Logarithmic energy values computed after converting an audio signal into Mel-frequency without application of the direct cosine transform, as proposed in [12,13], are used here for MFSCs. Several feature maps, e.g. after considering temporal derivatives of MFSCs, can be stacked on top of each other to form a 3-dimensional array, similarly to the RGB channels in image data.

This work considers several variants of short-term audio features, resulting in nine input feature maps in total:

1. Mel-frequency spectral coefficients ($MFSC$).
2. First temporal derivative of MFSC ($MFSC_{\Delta}$).
3. Second temporal derivative of MFSC ($MFSC_{\Delta\Delta}$).
4. Levinson-Durbin reflection coefficients (LRC) [14].
5. Vocal tract area coefficients ($VTAC$) [15].
6. Ratio of the area of the two successive vocal tract tubes calculated along the frames ($VTAC_R$) [15].
7. Classical spectrogram - frequency spectral coefficients (FSC).
8. Linear predictive coding coefficients ($LPCC$).
9. Mel-frequency cepstral coefficients ($MFCC$).

Figure 1 illustrates an example speech signal and input feature maps extracted from it. The number of coefficients (ticks on the vertical axis) considered was 80.

3.2 Convolutional Neural Network

The convolutional neural network (CNN) is a variant of standard neural network. Instead of fully connected layers CNN has architecture composed of consecutive pairs of *convolution* and *pooling* layers. CNN enables learning of local features and promotes weight sharing, where internal representations are a result of convolving the input with a filter mask. Each convolutional layer has a number

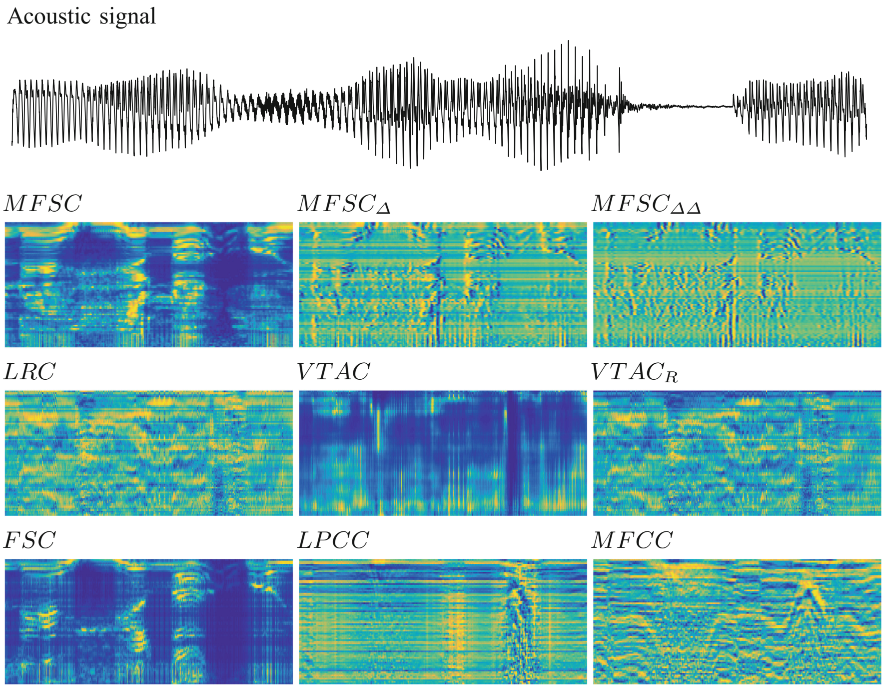


Fig. 1. An example of input feature maps extracted from speech signal.

of filter masks, which are learned during training, and application of convolution results in a new feature map, which is downsampled to a smaller size by using a pooling layer. Input feature maps in CNN are usually organized as a 3-dimensional array, where flat 2D planes in such an array are stacked input feature maps. For example, when CNN is applied on images, an array of 2D input feature maps corresponds to RGB channels. As shown in Fig. 2, every input feature map $\mathbf{O}_i(i = 1, ..., I)$, where I is the number of channels, is connected to further feature maps $\mathbf{Q}_j(j = 1, ..., J)$ through application of convolution. The i-th input feature map is connected to j-th feature map through convolution with a local weight matrix $\mathbf{w}_{ij}$, known as a filter mask. The filter mask is defined by a size $(m \times n)$, where (m) corresponds to a few spectrogram frequencies and (n) to a small temporal window. A non-input feature map $\mathbf{Q}_j$ is obtained using the convolution operation $*$:

$$\mathbf{Q}_j = \sigma\left(\sum_{i=1}^{I} \mathbf{O}_i * \mathbf{w}_{ij}\right) \qquad (j = 1, ..., J) \tag{1}$$

where $\mathbf{O}_i$ is the i-th input feature map, $\mathbf{w}_{ij}$ corresponds to a filter mask, σ is an activation function of the neural network. Therefore, CNN could be considered as a transformation of an input image through a chain of convolutions by the filter masks w learned from the data.

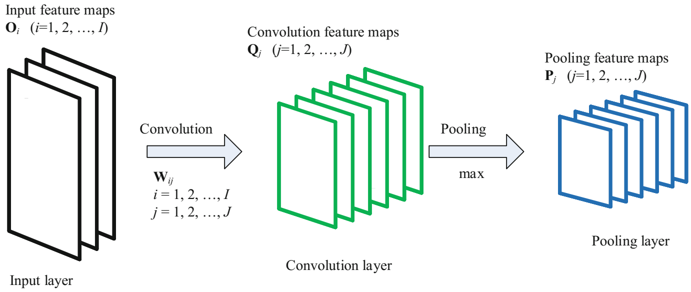


Fig. 2. A pair of *convolution* and *pooling* layers in the CNN architecture.

As shown in Fig. 2, the *max-pooling* operation is performed on each feature map, resulting from the convolution. The goal of *pooling* is to downsample the feature maps to smaller resolution. The max-pooling operation can be written as:

$$p_{ijk} = \max_{n=1,m=1}^{G} q_{i,(j-1)\times s+n,(k-1)\times s+m} \tag{2}$$

where G is the pooling size, s denotes the *shift size* (by how many pixels the pooling window is shifted), q_{ijk} is the jk-th element of i-th convolutional feature map Q.

After a pair (or several pairs) of *convolution-pooling* layers, the CNN is completed using a fully-connected dense layer, which uses feature maps from the last pair of *convolution-pooling* layers flattened into a one-dimensional vector. The output layer, connected to the dense layer, realizes the detection task by the *soft-max* of 2 neurons. An example of the CNN architecture, containing two pairs of *convolution-pooling* layers, is illustrated in Fig. 3.

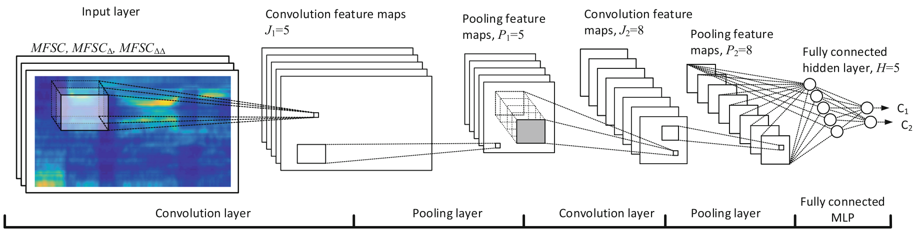


Fig. 3. An example of the CNN architecture, having *convolution-pooling* layers and a fully-connected multilayer perceptron (MLP) with 2 output neurons (predicted classes).

4 Experimental Investigations

A CNN with 4 pairs of *convolution-pooling* layers and a fully-connected dense layer with 2 *soft-max* neurons in the output (as shown in Fig. 3) was used for the experiments. The number of neurons in the dense layer (hidden layer of MLP before the *soft-max*) was 256, whereas other parameters of the CNN architecture are listed in Table 2. Filter masks of the first *convolution* layer had rectangular-shaped dimensions of 7×5, as recommended by [16] for spectrogram data, whereas filters in subsequent convolutions had traditional square-shaped dimensions of 3×3.

Table 2. Parameters for 4 pairs of the *convolution-pooling* layers.

Pair of layers	Convolution layer		Pooling layer
	# of maps	Filter size	Window size
I	96	7×5	3×3
II	256	3×3	3×3
III	384	3×3	3×3
IV	256	3×3	3×3

Length of speech recordings varies, therefore the size of the input feature maps is not constant with respect to the temporal axis. Input feature maps of fixed size were obtained using the following techniques: (a) cutting out signal fragments of 80 pixels in width at n random locations (F_{80}); (b) cutting out signal fragments of 120 pixels in width at n random locations (F_{120}); (c) re-sizing (squeezing or extending) a spectrogram image using the bi-cubic interpolation into a fixed width, corresponding to the mean width of all images (F_{mean}); (d) squeezing spectrogram image to the minimum width of all images (F_{min}). The number of random locations to sample from in the F_{80} and F_{120} cases was set to $n = 41$, which corresponds to the factor of data augmentation. Therefore, each subject, represented by a speech recording, has a single data example in the case of F_{mean} and F_{min}, but n examples in the case of F_{80} and F_{120}.

Detection performance was evaluated using the stratified 20-fold cross-validation. When performing data splits for the cross-validation, subject-level disjointedness was respected and all data examples for a subject (in F_{mean} and F_{min} cases) were either in a training or in a test fold. The equal error rate (EER), measured at the operating point of detector where sensitivity becomes equal specificity (or false alarm becomes equal miss rate) [17], was obtained for test data and is reported as a goodness-of-detection metric.

Results of initial experiments, not reported here, indicated improved detection performance when using all nine available input feature maps compared to a smaller combination or a single feature map. Detection results when using all nine feature maps are reported in Table 3. Interpolation by squeezing input maps to the smallest width (F_{min}) had slightly better performance than interpolation to the average width (F_{mean}). Nonetheless, the type of fragmentation did not affect the EER considerably. Meanwhile, the choice of sentence segment had a stronger influence on the EER varying from 20.3% for the ZILA segment to 29.5% for the TUREJO segment.

Table 3. Detection performance by EER (in %) for each type of fragmentation and sentence segment when using all input feature maps (shown in Fig. 1). Mean EER through all types of fragmentation is in the last line.

Type	TUREJO	SENELE	ZILA	OZELI	LE_O	RE_SE	TU_LE	ZI_LI	SENTENCE
F_{mean}	30.5	***23.1***	25.1	26.4	27.0	25.3	24.6	25.9	23.4
F_{min}	29.5	23.4	21.9	25.6	24.6	26.7	23.1	24.7	***21.4***
F_{120}	30.6	23.2	***20.3***	24.0	24.6	28.7	26.1	24.7	24.6
F_{80}	34.0	23.8	***22.0***	24.0	24.7	28.3	27.6	25.6	24.9
Mean	31.2	23.4	22.4	25.0	25.2	27.2	25.4	25.2	23.6

Due to the fact that the dataset was expanded in the case of F_{80} and F_{120} types, the example-wise EER shown in Table 3 is not particularly informative subject-wise. To obtain the subject-wise EER, the output class probabilities of n examples from a single recording were fused by averaging. Results in the subject-wise EER form after such fusion are given in Table 4. We can notice lower EER values, especially for the longer segments. It is also worth mentioning that data augmentation resulting from the random sampling (F_{80} and F_{120}), after fusion tends to outperform the interpolation approaches (F_{mean} and F_{min}) irrespective of sentence segment used.

Table 4. Detection performance by EER (in %) after fusing subject-wise decisions from n examples through averaging of output class probabilities.

Type	TUREJO	SENELE	ZILA	OZELI	LE_O	RE_SE	TU_LE	ZI_LI	SENTENCE
F_{120}	27.3	18.7	18.1	22.2	20.4	25.8	20.9	20.5	***16.8***
F_{80}	35.0	19.3	20.2	21.3	18.8	22.2	21.2	20.3	***17.1***

Aiming to improve over the best detection result (EER of 16.8% when using the SENTECE segment and the F_{120} fragmentation type), decision-level fusion of all nine segments was considered. Fusion of decisions arising for each kind of sentence segment (and n examples in the F_{80} and F_{120} cases) was done by: (a) voting (*vote*), where the majority class is considered; (b) averaging output probabilities (*prob*); (c) weighted average of output probabilities (*prob_w*), where weights are set based on the accuracy obtained using that segment; (d) the random forest [18] classifier (*RF*). The results of the decision-level fusion of sentence segments are given in Table 5. We can again notice that interpolation deteriorates detection, but now re-sizing to the average width outperforms squeezing. Meanwhile, the lowest EER of 14.1% was achieved for the (b) fusion case when random sampling of short fragments (F_{80}) was used.

Table 5. Detection performance by EER (in %) for decision-level fusion of all sentence segments. Results are reported by fragmentation type and fusion variant.

Type	*vote*	*prob*	*prob_w*	*RF*
F_{mean}	18.4	18.2	***17.8***	21.0
F_{min}	***19.2***	20.0	19.8	20.1
F_{120}	15.8	14.9	***14.8***	16.8
F_{80}	15.0	***14.1***	14.3	14.4

5 Conclusions

This work investigated PD detection from a speech signal using convolutional neural networks. Spectrograms and several other types of short-term features were considered as stacked 2D input maps to the CNN. A speech recording was split into various sentence segments and influence of each segment to the PD detection performance was evaluated and compared to the decision-level fusion of all segments case. The detection performance measured in EER varied from 29.5% for the TUREJO segment to 20.3% for the ZILA segment. This indicates that some parts of speech recording are more effective for the PD detection task than others.

Interpolation of spectrogram to the fixed length could not outperform the case of using fragments of fixed length taken at random locations and resulted in worse performance when fusion was considered. Therefore, data augmentation, arising from sampling of fixed length fragments, can be considered as more beneficial for CNN than using interpolation and less data.

The best detection result, EER of 14.1%, was achieved when using the decision-level fusion of sentence segments, whereas the best result without splitting a sentence into segments showed EER of 16.8%. Therefore, splitting a speech signal into several potentially overlapping segments and later combining decisions, obtained on each segment as well as a full-length signal, helps to improve PD detection. In this work splitting was done manually, but an automatic way of segmentation for text-dependent recordings should be devised in the future.

Acknowledgements. Funding for this work was provided by a grant (No. MIP-075/2015) from the Research Council of Lithuania. The dataset was collected by the Department of Otorhinolaryngology at Lithuanian University of Health Sciences.

References

1. de Rijk, M., Launer, L., Berger, K., Breteler, M., Dartigues, J., Baldereschi, M., Fratiglioni, L., Lobo, A., Martinez-Lage, J., Trenkwalder, C., Hofman, A.: Prevalence of Parkinson's disease in Europe: a collaborative study of population-based cohorts. Neurologic diseases in the elderly research group. Neurology **54**(11 Suppl 5), S21–S23 (2016)

2. Orozco-Arroyave, J.R., Hönig, F., Arias-Londoño, J.D., Vargas-Bonilla, J.F., Daqrouq, K., Skodda, S., Rusz, J., Nöth, E.: Automatic detection of Parkinson's disease in running speech spoken in three different languages. J. Acoust. Soc. Am. **139**(1), 481–500 (2016)
3. Hinton, G., Deng, L., Yu, D., Dahl, G.E., Mohamed, A.R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T.N., Kingsbury, B.: Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process. Mag. **29**(6), 82–97 (2012)
4. Abdel-Hamid, O., Mohamed, A.R., Jiang, H., Deng, L., Penn, G., Yu, D.: Convolutional neural networks for speech recognition. IEEE/ACM Trans. Audio Speech Lang. Process. **22**(10), 1533–1545 (2014)
5. Sainath, T.N., Kingsbury, B., Saon, G., Soltau, H., Mohamed, A.R., Dahl, G., Ramabhadran, B.: Deep convolutional neural networks for large-scale speech tasks. Neural Netw. **64**, 39–48 (2015). Special Issue on "Deep Learning of Representations"
6. Zhang, H., McLoughlin, I., Song, Y.: Robust sound event recognition using convolutional neural networks. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 559–563, April 2015
7. Thomas, S., Ganapathy, S., Saon, G., Soltau, H.: Analyzing convolutional neural networks for speech activity detection in mismatched acoustic conditions. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2519–2523, May 2014
8. Han, Y., Lee, K.: Acoustic scene classification using convolutional neural network and multiple-width frequency-delta data augmentation. Computing Research Repository (CoRR) arXiv:1607.02383 (2016)
9. Dennis, J., Tran, H.D., Li, H.: Spectrogram image feature for sound event classification in mismatched conditions. IEEE Signal Process. Lett. **18**(2), 130–133 (2011)
10. Deng, L., Abdel-Hamid, O., Yu, D.: A deep convolutional neural network using heterogeneous pooling for trading acoustic invariance with phonetic confusion. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6669–6673, May 2013
11. Adi, Y., Keshet, J., Goldrick, M.: Vowel duration measurement using deep neural networks. In: 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1–6, September 2015
12. Godino-Llorente, J.I., Gomez-Vilda, P.: Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. IEEE Trans. Biomed. Eng. **51**(2), 380–384 (2004)
13. Dibazar, A.A., Narayanan, S., Berger, T.W.: Feature analysis for automatic detection of pathological speech. In: Proceedings of the 2th Joint EMBS/BMES Conference, Houston, USA, pp. 182–183 (2002)
14. Verikas, A., Gelzinis, A., Vaiciukynas, E., Bacauskiene, M., Minelga, J., Hållander, M., Uloza, V., Padervinskis, E.: Data dependent random forest applied to screening for laryngeal disorders through analysis of sustained phonation: acoustic versus contact microphone. Med. Eng. Phys. **37**(2), 210–218 (2015)
15. Muhammad, G.: Voice pathology detection using vocal tract area. In: 2013 European Modelling Symposium, pp. 164–168, November 2013

16. Hrúz, M., Kunešová, M.: Convolutional neural network in the task of speaker change detection. In: Ronzhin, A., Potapova, R., Németh, G. (eds.) SPECOM 2016. LNCS (LNAI), vol. 9811, pp. 191–198. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-43958-7_22
17. Faundez-Zanuy, M., Monte-Moreno, E.: State-of-the-art in speaker recognition. IEEE Aerosp. Electron. Syst. Mag. **20**(5), 7–12 (2005)
18. Breiman, L.: Random forests. Mach. Learn. **45**, 5–32 (2001)

How to Increase Boys' Engagement in Reading Mandatory Poems in the Gymnasium: Homer's "The Odyssey" as Transmedia Storytelling with the Cyclopeia Narrative as a Computer Game

Ademir Pasalic, Nikolaj Hollænder Andersen, Christopher Schinkel Carlsen, Emil Åberg Karlsson, Markus Berthold, and Thomas Bjørner(✉)

Aalborg University, Department of Architecture, Design and Media Technology, A.C. Meyersvænge 15, 2450 Copenhagen SV, Denmark
tbj@create.aau.dk

Abstract. This paper outlines how a computer game can be used within transmedia storytelling to engage boys in the Danish gymnasium to read the epic poem "The Odyssey" by Homer. The study is based on a formative evaluation with questionnaires, observations, data logging, and interviews. Interviews with classical civilization teachers were further conducted, both in the initial design stage and after participant tests. Through all stages of the transmedia storytelling, it was revealed that curiosity was the main reason behind the engagement. The aesthetics worked out well in the game, and the story was presented in a format that the boys could comprehend and had the potential to increase their engagement in reading the poem. Different elements were implemented in the program code to maintain match in game flow for the individual player. However, this study can also emphasize how well-designed game mechanics are of ultimate importance, as well as involving teachers throughout all stages in the iterative process, as they are key figures for real implementation and acceptance.

Keywords: Game-based learning · Transmedia · Serious games

1 Introduction

Serious games within game-based learning are sometimes focused too much on the game itself. However, in terms of the overall learning goals, curriculum, content, and teachers' acceptance, one perspective could be to design the game-based learning as a transmedia element, where the game is not a standalone medium but incorporated within a broader context with other types of media. In this study, we will outline an approach for engaging boys in particular in the gymnasium, shifting between two media: a poem and a computer game. The idea is to create an engaging storytelling experience, meaning that the target group loses consciousness of the medium and neither sees the poem nor the game, but only the power of the story itself [1]. The

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 216–225, 2018.
https://doi.org/10.1007/978-3-319-76111-4_22

Danish Gymnasium offers a 3-year general academically oriented upper secondary program. Students are normally between the ages of 15 and 20. The Gymnasium is just one of several secondary education systems (others have specialized programs, such as commercial or technical) in Denmark, which all qualify students for admission to higher education. However, there are some common challenges for boys in particular in the Danish education system [2, 3]. In the gymnasium, boys struggle more with engagement in reading compared to girls. In general, there is a huge gender gap in reading, with 53% of boys who rarely or never read, compared to 29% of girls [14]. The difference in grades between boys and girls in the Danish gymnasium has also increased in favor of the girls [2, 3]. The gymnasium attracts more girls than boys, with 62% girls and 38% boys [3], and boys have a higher dropout rate (20%) than girls (15%) [3]. Boys lagging behind in education is not only a Danish challenge but also reported in other countries [4–6]. The reasons behind this gender gap are rather complex and multifaceted, but are, for example, described within genetic differences in ability, gendered learner identities and behavior, expectations of male and female identities, and the feminization of the education sector [6–8]. Common criticism is that schooling has favored female learners by using left-brain processes of fine motor skills, sequence, letters and words, and modes of teaching including "sit down and listen" [6, 9]. However, there are different initiatives for decreasing the gender gap in education, such as multifold initiatives in serious games.

Serious games are defined and used in divergent ways [10] but are commonly defined as games that have a focus on interactive computer-based software, which are developed with the intention to be more than entertainment [10, 12]. Serious games fit some of the general learning requirements like a glove, as they can provide a bridge to engage reluctant learners [11, 12]. Serious games offer intrinsically motivating game play and immediate feedback in the game environment, and the content often includes learning opportunities [10]. However, there are common examples where serious games fail within educational contexts, which could be due to poor design and a lack of knowledge about the target group. Failure could also be caused by, for example, the games not being very specific in terms of covering the curriculum, the games not simultaneously motivating and delivering a high quality of learning, the teachers not knowing how the game works or lacking IT skills, not having support to get started, and the preparation for playing the games taking too long [13].

This study is focusing on the example of the epic poem "The Odyssey" by Homer. "The Odyssey" is a part of the curriculum in the Danish gymnasium class "classical civilization." Teachers already use several alternative ways to teach classical civilization and "The Odyssey," such as watching movies, in-class role-play, drawings, YouTube clips, and going to the theater. However, there are no examples of using computer games within classical civilization in the Danish Gymnasium. The research question within this study is: How can a serious game be used in the Danish Gymnasium as transmedia storytelling in combination with Homer's "The Odyssey" to increase boys' engagement in reading the epic poem?

2 Engaging Transmedia Storytelling

The typology of transmedia storytelling used in this study is what Ryan defines as a one-world, many texts relation [15]. We are going to recreate an interpretation of the story of Odysseus within the original story. This is what computer games can achieve compared to the poem. The idea is to use interactive involvement to achieve knowledge within the rather complex texts of “The Odyssey,” but without the cognitive load of decoding text and making meaning from a bottom-up processing [11]. By using top-down processes [11] (by setting the main character as the player), it is possible within the game to create a world revolving around the player. This makes students free to explore or proceed at their own pace, with possibilities for allowing students to create an identity and engaging learners by involving them personally in the story [11, 16, 17]. Instead of reading about Odysseus (the valiant) defeating Polyphemus (the man-eating giant), the players are the ones defeating him. We will use a 3rd-person view to represent Odysseus, and we want the players to role-play and figure out (just as Odysseus did) how to defeat Polyphemus. We want the players to gain knowledge about Odysseus’s arrival to the island of Cyclops (Sicily), the transition into the cave where the dangerous Polyphemus lives, and how Odysseus manages to escape by getting Polyphemus drunk.

Based on the transmedia storytelling aspects in combination with user engagement theory [18–20] and player types [21], we have compiled an approach that aims to engage the boys in the activity of reading by a “game reward” when reaching book IX (the Cyclopeia narrative) in “The Odyssey.” After playing the game, the boys should be engaged to stay within the story world and continue reading with further discussions of the poem and related issues within classical civilization.

1. Point of engagement: The story has a point of engagement [19] before the student starts reading the epic poem so he knows there is something engaging to expect later on in the text reading. This is achieved by a teaser poster located either in the classroom or placed in the epic poem before the first page. The aesthetics from the computer game let the participant know that there is a game to play when reaching book IX.
2. Reading parts of the “The Odyssey,” book I to VIII.
3. Game prologue: The game prologue will have the attributes of aesthetics, motivation, and a clear goal for the game [19, 20]. The goal is to defeat Polyphemus and escape the cave. The prologue will provide the controls for the actual game. The objectives are to make the player understand that the controls are the keys WASD (to move around), the mouse (to look around), the space bar (to interact), and to invite the player to start playing (“pick up the wine by the barrels” and enter the cave).
4. Period of sustained engagement in game: For maintaining concentration and engagement throughout the game, relevant difficulty matches according to the player’s skill level (game flow channel) [22] were implemented. The game adapted the difficulty to the players’ skill level by measuring the time spent on each task and the performance in terms of amount of mistakes they made and amount of times they died. The game had four stages and evaluated the players’ performance for

each of these stages. Based on pretests, 20 s was an average time to complete the tutorial level, and it was rare that a test session lasted longer than 50 s. Incorrect key presses had a highest value of 10, and the mode was 3. With these parameters, the system normalized a value between 0 and 1 that represents both the duration in the level and the incorrect key presses. The value was stored in the file "PlayerPrefs" instead of the memory for later use in the game (e.g., if the player died in the game, the difficulty level could revert to the previous setting). Inserted below is a code example of one of the implementations for adjusting the difficulty level during the game play:

```
public void CalculateDifficulty() {
      float diff;
      if (timeOfEnd > 20)
      diff = (timeOfEnd - 20) / 30;
      else
      diff = 0;
      gameDifficulty /= 10;
      gameDifficulty += diff;
      gameDifficulty = Mathf.Clamp(gameDifficulty, 0, 1);
  PlayerPrefs.SetFloat("GameDifficulty," gameDifficulty);
      13 }
```

At this point in the game, the main level was loading, and Polyphemus chased the player. Polyphemus' speed was adjusted based on the player's previous difficulty score. A new timer began to calculate the amount of time that the player used to advance in the game. When the player completed the first objective in the main level, a new calculation was applied in order to adjust the game difficulty. The new difficulty was based on the duration of game play along with the amount of times that the player died and had to re-spawn.

5. Game epilogue and continued book reading: The game epilogue adds positive aspects to the story, which should make the player continue reading the poem. The game epilogue will be a cut scene, where the following is going to be explained: A. Odysseus escaped with his men. B. Odysseus is cocky and says his real name, displaying hubris. C. Polyphemus prays that his father, Poseidon, avenges him.

3 Methods

The method within this study consisted of a formative evaluation [24]. Within the formative evaluation, questionnaires, observations, data logging, and interviews were used to evaluate boys' engagement in reading "The Odyssey" through transmedia storytelling, including the game as book IX. The data collection took place in May and June 2017 at Aalborg University in Copenhagen, with 37 participants. An online asynchronous interview with one classical civilization teacher before game design and one interview with two classical civilization teachers after game design and participant test were further carried out.

3.1 Formative Evaluation

The formative evaluation was conducted using a seven-step strategy involving the following methods:

Step 1: Introduction and consent form. The participants were informed of the study purpose, that they were free to drop out at any time, and about anonymity.
Step 2: Participants were introduced to a short pre-experience questionnaire with questions regarding basic information (age, hours spent on computer gaming per week) and the initial engagement level for "The Odyssey" based on an NRS (Numerical Rating Scale, 0–10, with 0 = do not at all want to continue, and 10 = really want to continue).
Step 3: A teaser poster was introduced as an invitation to "The Odyssey" book IX and the game. Shortly after, participants filled in a questionnaire asking about two factors: how engaged they were to continue (same NRS scale as in step 2) and what characterized their engagement. Ten different cards with different engagement characteristics were presented for the participants: Fun, Commitment, Adventure, Growth, Change, Curiosity, Knowledge, Aesthetics, Satisfaction, and Challenge. Below each of these words, there was a short description to explain the word stated. For example, below Aesthetics was the description "something pleasing to the eye".
Step 4: Data logging was performed with the following information during the play session: A. Overall time spent. B. Amount of wrong key presses. C. Number of deaths.
Step 5: One observer and note taker was responsible to make non-participant observations about everything concerning the test participants' state of mind, as well as potential struggles in the game. Moreover, it was also the note taker's responsibility to note if the test participants had any comments or feedback on elements in the game and/or the overall experience. The note taker was positioned in the background, diagonally behind the test participants to be able to see the screen with the running game and note down observations with ease while the participants remained feeling comfortable.
Step 6: Short questionnaires were administered at different stages throughout the transmedia storytelling: A. When introduced to the first shift in media (the beginning of book IX). B. After the first stage had been completed. C. After playing the game. D. After the final objective and the epilogue within the epic poem were completed.
Step 7: A semi-structured interview was conducted, where participants were asked about their overall experience in the game, suggestions for improvements, and opinions about the transmedia shifts, as well as if they felt engaged in reading "The Odyssey" after playing.

3.2 Interviews with Classical Civilization Teachers

We conducted two interviews with classical civilization teachers. The first interview was an early stage interview performed as an asynchronous structured online interview. The purpose of the first interview was to gather initial game design ideas. The second interview was performed as in-depth interview with two teachers after game development and the participant test. The purpose was to get the teachers' opinions about the developed game and if it could be implemented in their classical civilization classes to engage boys in the reading. The second interview took place at a Danish Gymnasium in June 2017, following a semi-structured interview guide and analyzed as meaning condensation [23].

4 Design and Implementation

The game was implemented in the Unity3D game engine. It is playable on both PC and Mac. The style of the game is low polygon, with the intention of aesthetically sustaining or increasing the player's engagement and initial attraction, while also requiring relatively little to create. Several camera filters, particle systems, and assets were added to further increase the aesthetic level. They are accompanied by animations to fit each character with their movement, actions, and audio lines.

The prologue starts with the camera in an overview look so the players can quickly familiarize themselves with the scenery. Then, the camera slowly transitions down and toward Odysseus and stops when it reaches a 3rd-person perspective. This makes sure that the players know that they are in control of the main character, as it is necessary to tell the players that they are the ones playing and controlling Odysseus. This was coded through a Vector3.Lerp to interpolate with steps from the starting position and rotation, toward the ending position and rotation. This runs in a loop and progresses with T from 0 to 1. The program code is below, and Fig. 1 illustrates the prologue. On the first line,

Fig. 1. Screen shot from the initial view in the game

the size of each step is specified. On lines 2–5, the starting point and ending point are defined. On lines 8–13, the process is applied.

```
1 float step = 0.002f;
2 Vector3 fromAngle = mainCamera.transform.eulerAngles;
3 Vector3 fromPos = mainCamera.transform.position;
4 Vector3 targetRot = mainCamera.transform.eulerAngles;
5 Vector3 targetPos = mainCamera.transform.position -
    new Vector3(14.46f, 5.7f, 0f);
6 float t = 0;
7
8 while (t <= 1) {
9 t += step;
10 mainCamera.transform.eulerAngles =
    Vector3.Lerp(fromAngle,targetRot, t);
11 mainCamera.transform.position =
  Vector3.Lerp(fromPos, targetPos, t);
12 yield return null;
13 }
```

5 Findings

5.1 Engagement

There was increased engagement until the completion of the tutorial (mean = 7.87). After the encounter with Polyphemus, there was a significant drop in the engagement (mean = 6.96). The reduced engagement suggests that there were some problems with that level in the game (Fig. 2).

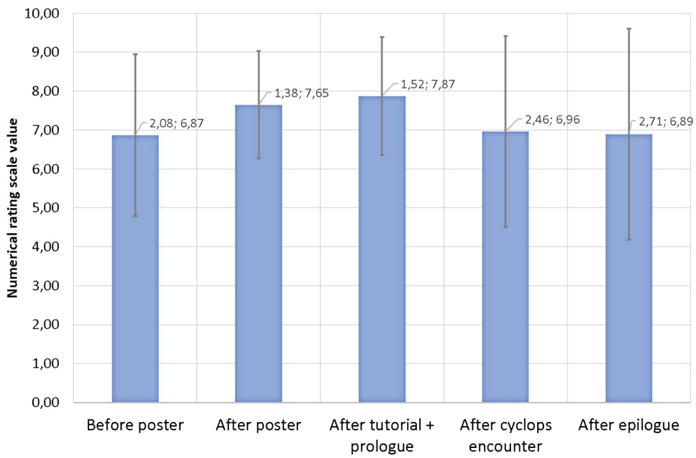


Fig. 2. Engagement level through different stages. NRS-scale 0–10, n = 37.

Through all stages of the test, curiosity (based on card selection) was the main reason behind the engagement to continue, especially after the poster, as well as after the tutorial and prologue stages. The curiosity card was chosen 18 times after the poster stage and 19 times after the tutorial stage. After the encounter with Polyphemus, only 8 participants selected the curiosity card as their main engagement factor. However, the number of curiosity card choices increased to 13 after the epilogue. This suggests that more participants became curious to continue after watching the epilogue, even though there were some problems in the level with Polyphemus. The drop in the curiosity engagement after the encounter with Polyphemus indicates that the participants lost some engagement after completing it and tended to continue the experience for reasons other than curiosity. Instead, they wanted to continue because of fun (6 times) or commitment (4 times). The general engagement level did not increase after the epilogue stage (mean = 6.89).

In the interviews with classical civilization teachers after game development, it was rather clear that they believed the game could have potential to engage boys more in the reading:

> "The boys are not good readers. They bridle at the poem format. With this game I am sure it could engage the boys as they feel familiar in gaming… To have challenges within the game media" (Kate: Teacher 1). "Also because they can participate in a rather interactive way" (Ann: Teacher 2).

5.2 Effectiveness

Of the observations noted in the Polyphemus encounter, 55% revolve around misunderstandings of game mechanics, while only 21% regard the mechanics as being understood. This implies that the game fails to provide the player with clear instructions, resulting in confusion. Furthermore, 21% revolve around technical problems such as bugs or program crashes. These problems could potentially have broken the engagement of the players, leading to a lower desire to continue with the game. The objective in the game that had the highest amount of misunderstandings was the stage where the player needed to help his men escape the cave. In some cases, the participants succeeded in saving some men, but usually, this was a result of pressing the interact button randomly. In the first objective, where the player would have to place the wine, most players used a lot of trial and error before figuring out the solution.

In the interviews with the participants, they also frequently mentioned the lack of guidance and the poor game mechanics. Mostly, they seemed confused when the game ended and they had yet to figure out why. Participants described the final objective in the cave as "I was frustrated that I could not figure out how to save the men", "The worst thing was the men escaping the cave", and "There was not enough time to save the soldiers". In some cases, the layout of the controls seemed unintuitive, since some expected the character to jump when using the space bar, rather than interacting. Others were confused with the functionality of interaction, stating, "It confused me that I could spam the interact button". Apart from what went particularly bad, we asked the participants, "What worked particularly well in the game?" The majority of the answers were that the aesthetics work well in the game, followed by the game play. In terms of the aesthetics, the art style seemed to fit the story well, as participants described it as

“Simple but meaningful” and “Nice and colorful” and stated, “The feel and tone of the environment was good”.

6 Conclusion

Boys have a higher dropout rate in the Danish Gymnasium and are lagging behind in terms of intake and marks. The reason behind this gap is rather complex and multi-faceted, but one perspective could be to implement game-based learning as a trans-media element within some of the very specific content in the classes, and where boys in particular are struggling. This could also help the game design to be very specific in terms of covering the curriculum. This would require the game to be engaging and deliver a high quality of learning, as well as not require too long to prepare for. However, for game-based learning, it is crucial that both the target group and the teachers are involved. The teachers’ knowledge is very important both in terms of how they could use the game-based learning in their classes, as well as making sure that the game content is professionally correct. In this study, we proposed a setup for increasing boys’ engagement in reading Homer’s “The Odyssey” through transmedia storytelling with the Cyclopeia narrative (book IX) as a computer game. Classical civilization teachers liked the idea behind this study and could see potential for real implementation in their classes to engage the boys in particular in the poem format. However, much further collaboration with the teachers through the iterative process would have improved the setup. We found that the participants had increased engagement after looking at a poster as the very first thing before reading. This suggests that an invitation for later gaming could work as an engagement factor for reading. Through all stages of the transmedia storytelling, it was revealed that curiosity was the main reason behind the motivation to continue. It was also revealed that the aesthetics worked out very well in the game, and the story was presented in a format that the boys could comprehend and had potential to increase the engagement in reading the poem. However, the game needed more clear objectives and some adjustments in terms of controls and further guidance, as there were too many misunderstandings within the game mechanics.

References

1. Murray, J.H.: Hamlet on the Holodeck. Free Press, New York (1997)
2. DEA Memorandum: Gymnasiet taber drengene [The Gymnasium are loosing the boys] (2014). https://www.dea.nu/sites/dea.nu/files/dea_notat_gymnasiet_taber_drengene.pdf
3. Hutters, C., Nielsen, M.L., Görlich, A.: Drenge og piger på ungdomsuddannelserne [Boys and Girls in Upper Secondary Education]. CEFU, Report (2013)
4. Smith, E.: Key Issues in Education and Social Justice. Sage, London (2012)
5. Clark, C., Burke, D.: Boys’ Reading Commission: A Review of the Existing Research to Underpin the Commission. National Literacy Trust, London (2012)
6. Hamilton, P.L., Jones, L.: Illuminating the ‘boy problem’ from children’s and teachers’ perspectives: a pilot study. Education 3–13 **44**(3), 241–254 (2016)
7. Sundaram, V.: Gender and Education. In: Arthur, J., Davies, I. (eds.) Education Studies Textbook, pp. 50–60. Routledge, Oxon (2010)

8. Legewiea, J., DiPretea, T.: School context and the gender gap in educational achievement. Am. Sociol. Rev. **77**(3), 463–485 (2012)
9. Featherstone, S., Bayley, R.: The Cleverness of Boys. A&C Black, London (2010)
10. Ritterfeld, U., Cody, M., Vorderer, P.: Introduction. In: Ritterfeld, U., Cody, M., Vorderer, P. (eds.) Serious Games: Mechanics and Effects, pp. 50–60. Routledge, New York (2009)
11. Dubbles, B.: Video games, reading, and transmedial comprehension. In: Handbook of Research on Effective Electronic Gaming in Education [Internet]. IGI Global. http://dx.doi.org/10.4018/9781599048086.ch015
12. Young, M.F., Slota, S.T., Travis, R., Choi, B.: Game narrative, interactive fiction, and storytelling. In: Video Games and Creativity, pp. 199–222. https://doi.org/10.1016/b978-0-12-801462-2.00010-2 (2015)
13. Bjoerner, T., Hansen, C.B.S.: Designing an educational game: design principles from a holistic perspective. Int. J. Learn. **17**(10), 279–289 (2010)
14. Kulturministeriet [Ministry of Culture Denmark]: Faktaark om produktion og læsning af litteratur [Fact sheet about production and reading of literature]. Report (2013)
15. Ryan, M.-L.: Transmedial storytelling and transfictionality. Poet. Today **34**(3), 361–388 (2013)
16. Gee, P.: Learning by design: good video games as learning machines. E-Learning **2**(1), 5 (2005)
17. Raybourn, E.M.: A new paradigm for serious games: transmedia learning for more effective training and education. J. Comput. Sci. **5**(3), 471–481 (2014)
18. Schønau-Fog, H., Bjørner, T.: "Sure, I would like to continue": a method for mapping the experience of engagement in video games. Bull. Sci. Technol. Soc. **32**(5), 405–412 (2012)
19. O'Brien, H.L., Toms, E.G.: The development and evaluation of a survey to measure user engagement. J. Am. Soc. Inf. Sci. Technol. **61**(1), 50–69 (2009)
20. Sweetser, P., Wyeth, P.: Gameflow: a model for evaluating player enjoyment in games. Comput. Entertain. **3**(3), 1–24 (2005)
21. Fullerton, T., Swain, C.: The Structure of Games. In: Game Design Workshop, pp. 26–48 (2008)
22. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper Perennial, New York (1990)
23. Bjørner, T.: Data analysis and findings. In: Bjørner, T. (ed.) Qualitative Methods for Consumer Research. Hans Reitzels Forlag, Denmark (2015)
24. Bjørner, T.: Data collection. In: Bjørner, T. (ed.) Qualitative Methods for Consumer Research. Hans Reitzels Forlag, Denmark (2015)

Analysis of Users Behaviour from a Movie Preferences Perspective

Andrea De Salve[1,2], Barbara Guidi[1(✉)], and Laura Ricci[1]

[1] Department of Computer Science, University of Pisa, Pisa, Italy
{desalve,guidi,ricci}@di.unipi.it
[2] IIT-CNR, via G. Moruzzi 1, 56124 Pisa, Italy

Abstract. Despite their tremendous popularity, Online Social Networks (OSNs) have several issues related to the privacy of social users. These issues have motivated researchers to develop OSN services that take advantage of the decentralized platforms (such as P2P systems or opportunistic networks). Decentralized Online Social Networks (DOSNs) need specific approaches to manage the decentralization of social data. In particular, data availability is one of the main issues and current proposals exploit properties of the social relationships to manage it. At the best of our knowledge, there are no proposals which exploit similarity between users, expressed with the term homophily. Homophily has been well studied in existing sociology literature, however, it is not easily extensible in Online Social Networks due to the limitations of real datasets. In this paper, we propose a preliminary analysis of similarity of social profiles in term of movie preferences. Results reveal that user's friends are characterized by a different levels of similarity which can be exploited to propose solutions for the data availability problem.

Keywords: Decentralized Online Social Networks · Data availability Homophily

1 Introduction

Online Social Networks (OSNs) [3] are the most popular online applications that have changed the way of how people interact between them. During the years, OSNs have filled the human life and several personal data are shared on these platforms. A huge amount of privacy problems have been arisen, in particular in Facebook, which is the most well-known OSN. To overcome privacy issues, decentralized solutions have been proposed. Decentralized Online Social Networks (DOSNs) [7] are made up of a set of peers, such as a network of trusted servers, a P2P system or an opportunistic network, which collaborate with each other in order to provide the social services. A DOSN is able to face the main problems of centralized solutions (i.e. scalability, privacy, etc.). However, in a decentralized environment other important issues have to be managed, such as the data availability [16]. Indeed, DOSNs are built on top of the users' devices

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 226–236, 2018.
https://doi.org/10.1007/978-3-319-76111-4_23

and when a user goes offline, data which are stored on his device should be available. In this context, the knowledge of human behaviour can help to manage the data availability problem. In OSNs, users are characterized by a virtual profile which contains attributes representing their private information and interests. These information can be easily obtained and analysed to find similar behaviours or other important aspects.

As for instance, homophily [17] is the principle that a contact between similar people occurs more than among dissimilar people and it is useful to implement new replication strategies which take into account groups of users with specific interests to store data. Indeed, users tend to bond more with users who are similar and studies have shown that similarity between users is a good indicator of possible future interactions.

In addition, the similarity between users is an indicator of how fast the information spreads among users and it can be exploited either to limit or to speedup the dissemination of information to the users having common interests.

In this paper, we investigate the homophily by focusing on a user-centric point of view. We exploited the ego network model to structure the relations between the center user (named ego) and its friends (named alters) [13]. The relations between the ego and the alters are used to represent the movie preferences and the similarity between these preferences is evaluated. Furthermore, we evaluate the similarity between OSN users by exploiting the Dunbar's concept [12]. Our evaluation exploits a real Facebook dataset and all the studies concern movie genres expressed by likes to Facebook pages. Results shows that there is homophily between egos and its alters and, more in detail, the similarity is affected by the tie strength computed by considering the contact frequency between users in Facebook.

The rest of the paper is organized as follow. Section 2 describes the current proposals in both the DOSNs fields and the homophily in OSNs. In Sect. 3 we explain how use homophily in a DOSN. In Sect. 4 we introduce our Facebook dataset and in Sect. 5 we show our evaluation. Conclusions and Future Works are proposed in Sect. 6.

2 Related Work

In this section we describe current DOSN proposals and the main studies conducted in Online Social Networks to discover homophily between social users.

2.1 Current DOSNs

The first important proposal is Diaspora [1], which consists of a federated network of servers. On one hand, Diaspora represents the first example of a real distributed social network, on the other hand, its main drawback is the scalability, due to its architecture is not fully distributed. Other proposals exploit a fully distributed architecture, often implemented by P2P systems. In Safebook [6], data are stored in a particular social overlay named "Matryoshka", which are concentric rings of peers around each users peer. LifeSocial [14] is a DOSN focused on the privacy issue, where user information is stored by exploiting a Distribued Hash Table (DHT) and the OSN functionalities are realized by plug-ins. PeerSon [4] is a distributed infrastructure for social networks whose focus is related to security and privacy concerns. It proposes a two-tier architecture where the first tier is a DHT and the second tier consists of the nodes representing users. All users' content is encrypted. DiDuSoNet [15] is a DOSN based on the Dunbar's concept, where social data are stored only on trusted nodes.

2.2 Homophily in OSNs

Homophily [17] means that similar individuals associate with each other more often than others. Several studies have been performed, and a detailed summary is shown in [18]. However, to the best of our knowledge, there have been very few studies that involved analysis of OSNs to investigate the principle of homophily, probably for the lack of real OSNs datasets. Authors in [5] study the LiveJournal and Wikipedia data and used activities such as user edits to evaluate the similarity between individuals. In [2] authors proposed a systematic approach to study homophily concept on two online social media networks, *BlogCatalog* and *Last.fm*. Simsek and Jensen [19] have proposed a technique applied in distributed systems, for navigating networks by exploiting homophily. Results show that a simple product of degree and homophily measures can be quite effective in guiding local search. Finally, authors of [8] showed that the availability patterns of the egos and their alters increases when considering alters with a strong tie strength referred to the ego. In addition, alters of Dunbar's circles have similar temporal pattern.

3 Homophily as a Strategy for the DOSNs' Issues

DOSNs have several issues regarding the decentralization of social services. One of the main problem is the data availability, which occurs in distributed systems because data are stored among users and the online behaviour of them can affect the availability of social data. In DOSNs, replication is the most used technique to maximize data availability and it consists of storing copies of the same data on several devices and the users where data are allocated are named replica peers. The understanding of the user temporal behaviour is a crucial aspect for all those systems that rely on the users' resources for the daily operations, such as DOSNs.

Current approaches exploit properties of the social graph, such as the coverage of the social graph [9] and community structure [10,11] to face the problem of data availability. However, other important characteristics of users can be used to manage the problem of data availability. The homophily concept studied in this paper is suitable to be used for the data availability problem because we expect that users who have similar behaviour (in terms of movies preferences, music preferences, books preferences, etc.) are more interested to the same contents.

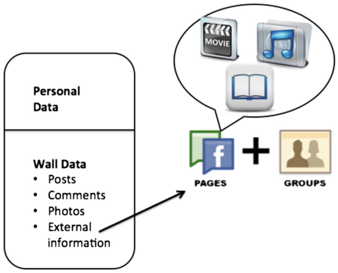


Fig. 1. User's profile overview.

As a result, the data availability strategies can be refined by allocating the data on the peers of the users having similar interests. User's profile contains several external information represented by pages and groups (Fig. 1) which can be used to measure similarity between users. Both a like operation on a Facebook page or the join to a group gives an important information to the characterization of a user.

4 The Facebook Dataset

Information about movie preferences of Facebook users have been gathered by a Facebook application, called SocialCircles![1], which exploits the Facebook API to retrieve social information about registered users. The application was able to retrieve information about the Ego Network of registered users. as explained in [8]. In detail, the application retrieved:

- Topology Information. We were able to obtain friends of registered users and the friendship relations existing between them.
- Profile Information. We downloaded profile information of registered users and their friends, such as complete name, birthday, sex, current location, hometown location, works, schools, user devices, movies, music, books, interests and languages.
- Interaction Information. We collected information about interactions between registered users and their friends, such as posts, comments, likes, tags and photo. Due to technical reasons (time needed to fetch all data and storage capacity), we restrict the interaction information retrieved up to 6 months prior to user application registration.

The dataset contains 337 complete Ego Networks, for a total of 144, 481 users (ego and their alters). The sample obtained from Facebook consist of 213 males and 115 females (while 9 users did not specify their gender) with age between 15 and 79, having different education, background and geographically location. We focus on the part of the profile which contains the movies that the profile

[1] http://social.di.unipi.it.

owner likes. About 77% and 58% of the registered users and the registered users' friends respectively, exposes preferences about movies. The registered users have, on average, 5 favorite movies while the registered users' friends have about 8 favorite movies. The most part of the registered users (90%) have a fraction of friends without favorite movies that do not exceed 0.6.

Our dataset contains 69, 519 movie titles. Each title is referred to a Facebook movie page that one of more our users (registered and their friends) have added to their profiles through the *like* button. Usually, titles includes several typos and they are written in several different languages. One of the main problem is that a huge amount of movie titles refers to same movies and titles are different for example, due to typos. A data preprocessing phase has been executed to obtain a refined movies dataset.

4.1 Data Preprocessing

The first step of the data preprocessing cleans dataset by excluding all titles which both are not referred to film and titles which contains no Latin characters. The total number of film after this step was 61, 918.

The second step concerns the issue of duplicate Facebook pages which refer to the same movie. In order to discover pages which refer to the same movie we use a set of similarity metrics based on string and the public movies dataset MovieDB[2]. The Movie Database (TMDb) is an open database of movies and television information concern movies, television shows, production companies, and individuals in the entertainment industry. The TMDb API is a RESTful web service to obtain movie information. All content and images on the site are contributed and maintained by users. The similarity measures used in the data preprocessing are combined to exploit the main properties of each of them. In particular, we use the following similarity metrics:

Cosine Similarity. It is a measure of similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them. It is a common vector based similarity measure.

Levenshtein Similarity. It is a metric for measuring the difference between two sequences. When we consider it as distance, it measures the minimum number of single-character edits required to change one word into the other.

Smith-Waterman Similarity. It is a well-known measure in the Edit-based similarity metrics. It is an algorithm finding the best local sequence alignment or using the longest common subsequence.

The similarity measures we considered take as input parameters the field *about* containing the information (or title) about the pages. The measures are combined as follow, in order to obtain a global similarity measure between two Facebook movie pages:

$$TitleSim = CosineSim * LevenshteinSIM * SmithWatermanSim \quad (1)$$

[2] https://www.themoviedb.org.

We computed the *TitleSim* measure for each pair of pages. In this way we were able to cluster similar pages related to the same movie, regardless of the typos or differences in the title among pages. Afterwards, clusters of pages have been used to interrogate the MovieDB dataset. When a query to MovieDB produces a reply, this reply is used inside the cluster to classify the pages which does not receive a response from MovieDB. Indeed, MovieDB is not able to reply in the case of typos but the *TitleSim* measure are able to understand which pages are similar. At the end of this phase, we obtain the final dataset used to execute our evaluation which contains 45,729 pages which have been enriched by attaching to each page the genre of the corresponding movie taken from MovieDB. The other pages (16,189) have not an associated genre, usually because MovieDB does not provide a reply or the obtained genre is empty.

5 Analysis of the Dataset

In this section we show the evaluation of the similarity between users as concerns the movie genres. After the preprocessing phase, explained in Sect. 4.1, we obtain a dataset of 45,792 movie pages. Each page has an associated genre which we use to categorize the behaviour of users.

5.1 Movie Genres

We identify 25 genres including the genre *unknown* which contains Facebook pages which are not classified by MovieDB.

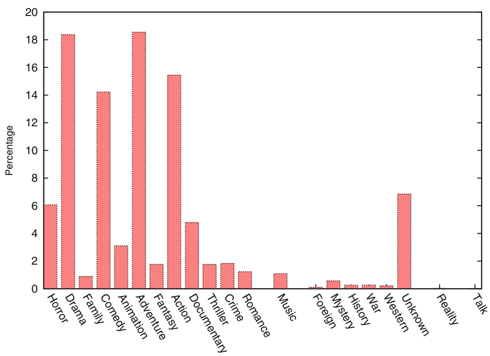


Fig. 2. Distribution of the genres.

Figure 2 shows the distribution of the genres. The majority of the pages are distributed among four principal genres: *Adventure*, *Drama*, *Action*, and *Comedy*. In particular, 20% of pages are classified as *Adventure*, about 18% of pages are classified as *Drama*, and about 15% of pages are classified as *Action* and *Comedy*.

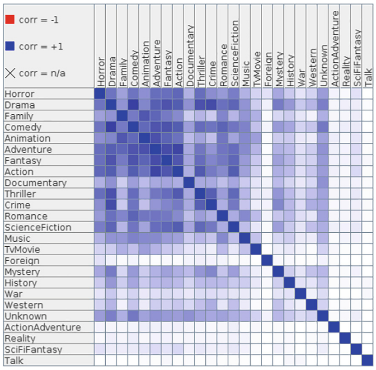


Fig. 3. Correlation matrix between preferences on different movie genres.

We represent the movie preferences of each user as an array of 25 items, corresponding to the distinct genres identified in the dataset. We investigate if users who like a specific genre also like other genres as well, by analyzing the correlation between genres liked by each user. Figure 3 shows the correlation matrix between movie genres liked by the users where color gradation represents the strength of the correlation (strong correlation corresponds to dark color). The matrix indicates the presence of a higher correlation between *Comedy* and *Drama*, but also *Crime* and *Thriller* with *Drama*. We have also a high correlation between *Action* and *Adventure*.

5.2 Evaluation of the Homophily in Ego Networks

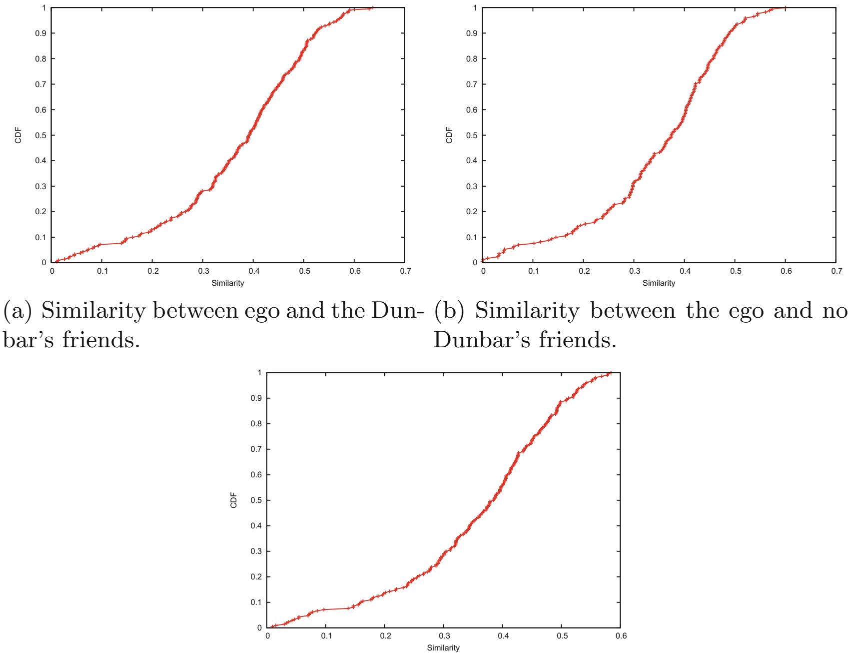
We start to study the homophily between a user with its friends by considering ego networks. We evaluate the similarity between an ego node and its alters by dividing them into three different sets:

- Dunbar's Friends. Dunbar [12] explains that human brain has cognitive limit to the number of people with whom one can maintain stable social relationships. He proposed that humans can comfortably maintain only 150 stable relationships. For these reason, relationships are classified according to the strength of the relation. The Dunbar friends of an ego are the friends with whom the ego has a stronger relation by taking into account the social interactions (posts, comments, etc.).
- No Dunbar's Friends. The Dunbar's Friends are limited to 150. The other alters contained in the ego that are not in the first 150 alters are automatically classified as no Dunbar's Friends.
- All alters. This set contains all the alters in the ego network of the users, i.e., the union of the Dunbar's Friends and the No Dunbar's Friends.

Table 1 reports some statistics about the ego networks considered in our experiments. Ego networks have on average 320 alters and the higher Standard

Table 1. Statistics of the ego networks.

Measure	Value
Number of ego networks	229
Min ego network size	28
Max ego network size	1394
Mean ego network size	320.707
St. deviation ego net. size	227.516



(a) Similarity between ego and the Dunbar's friends.

(b) Similarity between the ego and no Dunbar's friends.

(c) Cosine similarity between the ego and all its friends.

Fig. 4. Similarity between users by considering the movie preferences

Deviation suggests more variation in the number of alters in the ego networks. Indeed, the smallest ego network has 28 alters, instead the size of the largest one is 1,394.

We evaluate the similarity between the movie preferences of an ego and all its alters by using the Cosine similarity. Figure 4(c) shows the Cumulative Frequency Distribution (CDF) of the cosine similarity. About the 50% of egos have a similarity with its friends less than 0.4. However, more than the half of ego nodes show a high similarity (between 0.4 and 0.6). This means that the movie preference of an ego are similar to the half of its friends, by suggesting the presence of a sort of influence. We investigate in more details the similarity between the ego and the different sets of alters of the ego network. Figure 4(a) shows the CDF of the similarity between the egos and their Dunbar's friends while Fig. 4(b) shows the similarity between the egos and their No Dunbar's friends. The plots clearly indicate that users of the No Dunbar's friends expose lower similarity in terms of movie preferences. Indeed, about 50% of egos show a similarity lower than 0.37 with its No Dunbar's friends. Instead, the similarity between the movie preferences of the egos and those of their Dunbar's friends is slightly higher, i.e., half of the users show a similarity of about 0.4. This suggest that users that frequently interact with each other expose a higher level of similarity in terms of movie preference.

Consider that the half of ego nodes shows a high similarity, we decide to evaluate if there is a relation between the cosine similarity computed by considering all the friends and the cosine similarity computed on both Dunbar or no Dunbar friends. We evaluate the Pearson correlation shown in Table 2.

Table 2. Pearson correlation

	Dunbar's similarity	No Dunbar's similarity	Alters' similarity
Dunbar's similarity	1	0.308	0.994
No Dunbar's similarity	0.308	1	0.287
Alters' similarity	0.994	0.287	1

We can notice that there is a positive correlation between the similarity computed by considering all the friends and the cosine similarity computed by considering only Dunbar's friends, and a positive but more scattered correlation between the similarity computed by considering all friends and the similarity computed on the set of no Dunbar's friends. In addition, Table 2 indicates that users establish friendship relations with alters having similar interests.

6 Conclusion and Future Works

In this paper, we propose a preliminary analysis of the homophily in a real Facebook dataset to face the problem of data availability in DOSNs. We study the movie preferences similarity in ego networks by exploiting the MovieDB database to retrieve information about the genre of our Facebook pages. In terms of data availability, the homophily between ego and its alters, in particular with Dunbar's friends can be exploit to predict the content requests by analysing interactions and other properties of the social graph.

Results show that there is a high homophily between an ego node with its alters, in particular with its Dunbar's friends. We plan to investigate more in detail the similarity between users by analysing the temporal behaviour of users and we want to detail how Dunbar's friends impact on the similarity by analysing the Dunbar's circles. Moreover, we want investigate other feature of the social profile, such as music and/or book preferences to better understand the behaviour of users. Finally, we want propose a specific data availability strategy which takes into account the homophily studied in this paper.

References

1. Diaspora Website. https://diasporafoundation.org/
2. Bisgin, H., Agarwal, N., Xu, X.: Investigating homophily in online social networks. In: 2010 Web Intelligence and Intelligent Agent Technology (WI-IAT), vol. 1, pp. 533–536 (2010)
3. Boyd, D., Ellison, N.B.: Social network sites: definition, history, and scholarship. J. Comput.-Mediat. Commun. **13**(1), 210–230 (2007)
4. Buchegger, S., Schiberg, D., Vu, L.H., Datta, A.: PeerSoN: P2P social networking - early experiences and insights. In: SNS, pp. 46–52. ACM (2009)
5. Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., Suri, S.: Feedback effects between similarity and social influence in online communities. In: International Conference on Knowledge Discovery and Data Mining, KDD 2008, pp. 160–168 (2008)
6. Cutillo, L.A., Molva, R., Strufe, T.: Safebook: a privacy-preserving online social network leveraging on real-life trust. IEEE Commun. Mag. **47**(12), 94–101 (2009)
7. Datta, A., Buchegger, S., Vu, L.H., Strufe, T., Rzadca, K.: Decentralized online social networks. In: Furht, B. (ed.) Handbook of Social Network Technologies and Applications, pp. 349–378. Springer, Boston (2010). https://doi.org/10.1007/978-1-4419-7142-5_17
8. De Salve, A., Dondio, M., Guidi, B., Ricci, L.: The impact of users availability on on-line ego networks: a facebook analysis. Comput. Commun. **73**, 211–218 (2016)
9. De Salve, A., Guidi, B., Mori, P., Ricci, L.: Distributed coverage of ego networks in F2F online social networks. In: 2016 International IEEE Conferences ATC, pp. 423–431 (2016)
10. De Salve, A., Guidi, B., Ricci, L.: An analysis of ego network communities and temporal a affinity for online social networks. In: Gaggi, O., Manzoni, P., Palazzi, C., Bujari, A., Marquez-Barja, J.M. (eds.) GOODTECHS 2016. LNICST, vol. 195, pp. 135–144. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61949-1_15
11. De Salve, A., Guidi, B., Ricci, L.: Evaluation of structural and temporal properties of ego networks for data availability in DOSNs. Mob. Netw. Appl. 1–12 (2017)
12. Dunbar, R.I.M.: The social brain hypothesis. Evol. Anthropol.: Issues, News, Rev. **6**, 178–190 (1998)
13. Everett, M.G., Borgatti, S.P.: Ego network betweenness. Soc. Netw. **27**, 31–38 (2005)
14. Graffi, K., Gross, C., Stingl, D., Hartung, D., Kovacevic, A., Steinmetz, R.: LifeSocial. KOM: a secure and P2P-based solution for online social networks. In: IEEE CCNC (2011)
15. Guidi, B., Amft, T., De Salve, A., Graffi, K., Ricci, L.: Didusonet: a P2P architecture for distributed dunbar-based social networks. Peer-to-Peer Netw. Appl. **9**, 1177–1194 (2015)
16. Guidi, B., Conti, M., Ricci, L.: P2P architectures for distributed online social networks. In: 2013 International Conference on High Performance Computing and Simulation (HPCS), pp. 678–681. IEEE (2013)

17. Lazarsfeld, P.F., Merton, R.K.: Friendship as a social process: a substantive and methodological analysis. In: Freedom and Control in Modern Society, pp. 18–66, New York (1954)
18. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: homophily in social networks. Ann. Rev. Sociol. **27**(1), 415–444 (2001)
19. Şimşek, Ö., Jensen, D.: Navigating networks by using homophily and degree. Proc. Nat. Acad. Sci. **105**(35), 12758–12762 (2008)

Let's Cook: An Augmented Reality System Towards Developing Cooking Skills for Children with Cognitive Impairments

Eleni Papadaki[1(✉)], Stavroula Ntoa[1], Ilia Adami[1], and Constantine Stephanidis[1,2]

[1] Institute of Computer Science, Foundation for Research and Technology – Hellas (FORTH), N. Plastira 100, Vassilika Vouton, 700 13 Heraklion, Crete, Greece
{elpap,stant,iadami,cs}@ics.forth.gr
[2] Department of Computer Science, University of Crete, Heraklion, Greece

Abstract. Although activities of daily living are often difficult for individuals with cognitive impairments, their autonomy and independence can be fostered through interactive technologies. The use of traditional computer interfaces has however proved to be difficult for these users, bringing to the surface the need for novel interaction methods. This paper proposes *Let's Cook*, an innovative Augmented Reality game, designed to teach children with cognitive impairments how to prepare simple meals, following a playful approach. *Let's Cook* supports multimodal interaction techniques utilizing tangible objects on a tabletop surface, as well as multimedia output. Additionally, it can be personalized to accommodate the diverse needs of children with cognitive impairments by employing individual user profiling. The system is currently installed in the kitchen of the Rehabilitation Centre for Children with Disabilities in Heraklion, Crete where it was evaluated by the students.

Keywords: Augmented reality · Multimodal interaction · Adaptability Cognitive impairments · Meal preparation · Cooking

1 Introduction

Cognitive impairment is a broad term encompassing various intellectual or cognitive deficits and may vary from severe cognitive disability to mild impairment. A typical limitation among people with cognitive impairments is carrying out activities of daily living, such as cooking. To that end, assistive technologies have been developed, aiming to help individuals with cognitive disabilities to become more independent and enhance their quality of life [1]. The use of traditional computer interfaces has however proved to be difficult for these users [2], bringing to the surface the need for novel interaction methods.

Play is an important activity of life, in all its contemporary and digital forms, such as video games, applications on smartphones, interactive toys etc. [3]. Game-based methods, technologies, and concepts are employed by serious games with the aim to teach, exercise, and change users' behavior [4]. Playing and cognition are deeply

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 237–247, 2018.
https://doi.org/10.1007/978-3-319-76111-4_24

interconnected since learning "how to play" a computer game implies various cognitive skills [5]. Technologies like Augmented Reality (AR) have the ability to blend a fictional narrative with the real and familiar physical environment and to provide a high engaging learning environment [6].

Focusing on meal preparation, which is considered an important skill to teach to individuals with cognitive disabilities in order to increase their independence, this paper describes *Let's Cook*, an educational AR game, aiming to familiarize children with cognitive impairments with elementary meal preparation notions and simple recipes, in a playful way. Taking into account the need for novel interaction methods, the system supports multimodal input, utilizing tangible objects on a table-top surface and multimedia output available in textual, auditory and pictorial form.

Let's Cook is one of the five games comprising *CocinAR*, an educational system, which aims to instruct children: (i) which meals are appropriate for breakfast, lunch, and dinner, (ii) how to cook simple meals and (iii) fundamental rules of safety and hygiene that should be applied during the food preparation process [7]. Profiling functionality is supported, allowing the system to adapt to each individual student with respect to the preferred output modalities, providing at the same time teachers with statistic information on the progress of each student. *CocinAR* is installed in the kitchen of the Rehabilitation Centre (RC) for Children with Disabilities in Heraklion, Crete, where it is used by the educators in the context of their "Independent living" program, in which children are trained to execute simple recipes. Being *CocinAR*'s central game, *Let's Cook* incorporates the system's main functionality, allowing the user to "cook" a simple recipe in a virtual kitchen, following a step-by-step approach.

The remaining of this paper is structured as follows: Sect. 2 discusses related work, Sect. 3 provides an overview of the *CocinAR* system, Sect. 4 describes the *Let's Cook* game in detail, and Sect. 5 reports the results of an evaluation that was carried out with children with disabilities. Finally, the paper concludes with discussion and future work in Sect. 6.

2 Related Work

Activities of daily living (ADL) such as eating and taking a bath, are routine self-maintenance tasks. Regarding the promotion of ADL, technology has been claimed to improve performance of individuals with disabilities in comparison to more traditional means, such as plain pictorial or auditory prompts [8, 9]. Typical assistive technologies include software for reminding and prompting, task guidance, computer assisted learning and communication [1, 2, 10].

On the other hand, several serious games have emerged as a means of increasing the efficiency of learning methods. An AR application developed to assist children with autism spectrum disorders to recognize and acquire emotions is GameBook [11], which can be played on any mobile device, such as a tablet, a smartphone or a laptop. AR systems often involve interaction on an augmented table surface. For example, PAR [12] is a multiuser collaborative game for people with autism, which supports multi-touch tabletop interaction. However, touch-based interaction may prove to be challenging for very young children or children with physical and cognitive disabilities, on the grounds that their fine motor skills may not be sufficiently developed [14, 15]. Alternatives have appeared combining physical artifacts and computer-augmented surfaces, such as Beantable [17], a system developed to support preschool children's development and combines interaction via touch, augmented objects and physical cards.

Regarding the meal preparation process, portable devices have been proposed involving pictorial instructions, auditory and vibratory feedback [20], or multi-step tasks in a self-prompting approach using video, picture and auditory prompts [21]. Another step-by-step approach for guiding individuals with cognitive impairments through a food preparation recipe is the ARCoach system [13]. ARCoach uses computer vision to recognize the various ingredients on a table and the user's actions, so appropriate feedback is provided when there is an abnormality in a step execution. An assistive system which addresses people with higher brain dysfunction and aims to support the meal preparation process focuses on assisting their memory and planning ability through multimedia recipe presentation [22]. Targeted at meal preparation, with the aim to assist the better comprehension and use of cooking recipes *CounterActive*, is an interactive cookbook that is projected on the user's kitchen counter and can be interacted with via touch [23]. An interesting AR cooking system, which provides a first cooking experience with recipes to young children, employs a tabletop screen and a miniature kitchen to simulate cooking activities [16]. Interaction is accomplished via markers, which correspond to cooking elements on the miniature kitchen, and specific touch gestures on the tabletop display, which indicate the various commands (e.g., break egg, turn on fire etc.).

In summary, efforts in the domain of cultivating cooking skills in individuals with cognitive impairments have focused on multimedia information presentation, step-by-step presentation of the instructions, as well as simulation of the cooking activities or live assistance during the cooking process. *Let's Cook* addresses children who are in the process of learning how to cook and adopts the multimedia information presentation and step-by-step guidance approaches. Furthermore, it combines the benefits of immersive experience and blending familiar environments with virtual objects brought by AR, and the playful learning approach of serious games. Progress beyond the current approaches lies in the multimodal interaction it features, the adoption of a novel interaction style, the incorporation of symbol-based language, combined with the adaptability of the system to the needs of each individual child, through profiling mechanisms. The input and output mechanisms were carefully designed, aiming to achieve an intuitive and comfortable interaction with the system, while the various modalities can be provided concurrently, personalized to support different levels of cognition.

3 The *CocinAR* System

CocinAR has been implemented following an iterative process with the active participation of UX experts and educators of the RC, who provided the target user characteristics and the specific educational goals that the system should serve in the context of their "Independent Living" training program [7]. In summary, the system should be able to accommodate a large variety of skills, by: (i) supporting information presentation through text, audio, images and symbols, concurrently available if needed, (ii) allowing customization of font size, (iii) employing a minimalistic approach towards graphic design, (iv) providing context-sensitive help in each individual exercise, and (v) supporting two levels of content, a simple one for students with severe cognitive problems and a more advanced one, allowing teachers to switch to any of the two levels at runtime. In terms of functionality, *CocinAR* features four multiple choice games (meal appropriateness, collect the ingredients, collect the utensils, safety and hygiene rules) and *Let's Cook*, a cooking simulation game asking students to follow step-by-step instructions in order to prepare simple recipes [7].

CocinAR consists of a computer, a high-resolution projector, a simple wooden table, an infrared camera and a high-resolution camera. The system is designed to "camouflage" itself in a way that none of the equipment used is visible to the users. This is achieved by hiding the hardware equipment inside a kitchen board, leaving visible only the plain wooden table, with the aim to support an immersive user experience (Fig. 1). The interaction with the system is accomplished via simple printed cards, while any infrared device can be used to produce mouse-click equivalent functionality when pressed against the table. In order to facilitate an easier and more comfortable interaction, a custom infrared pointing device was designed for children with motor difficulties, which supports three different handles to accommodate different grip sizes [7].

(a) (b) (c)

Fig. 1. (a) Representation of the system setup (b) The actual setup at the RC (c) Physical cards and pointing device

4 The *Let's Cook* Game

Let's Cook simulates the cooking of a given recipe, following a step-by-step approach. All the recipes incorporated in the system were provided by the RC's educators who categorized them under three levels of difficulty. Based on the student's level, which is retrieved by the profiling module, *Let's Cook* suggests the appropriate recipe difficulty level. The recipe selection process is facilitated by statistics on each student's performance for each recipe as shown in Fig. 2, whereby the teacher can view at a glance the date that the recipe was last played and the score achieved.

Fig. 2. First level recipes with statistics

Once the recipe is selected, the game starts in a virtual kitchen, which visually resembles the actual RC kitchen where children are trained. The virtual kitchen contains points of interest (POIs), which are deliberately confined at the bottom half of the physical table, so that they are easily reachable by children. These interactive POIs, called slots, are the positions where the virtual ingredients and utensils appear and can be manipulated using the pointing devices.

The recipe steps appear sequentially at the top right corner of the virtual kitchen, as regular text or symbolic language, which replaces text words with explanatory images (Fig. 3). The instruction visualization depends on the student's current level and can be changed at runtime by the teacher through the game settings in the menu (target symbol at the top right corner of the virtual environment). Recipe step representation is accompanied by game-flow options addressing educators' requirements, allowing them to skip a step (e.g., create a sandwich without lettuce), to indicate that a step has been executed correctly by the student (e.g., when the teacher believes that the step was completed successfully in relation to the current session's goal), as well as to navigate back and forth in the recipe steps.

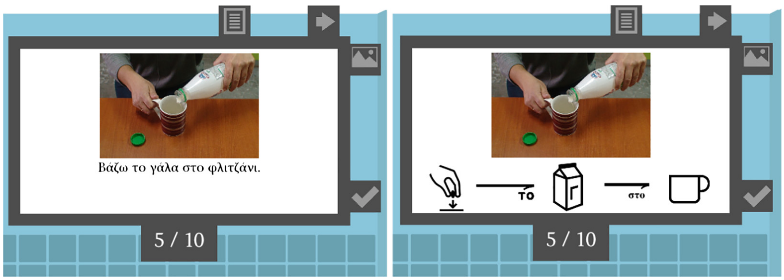


Fig. 3. (a) Question provided through text (b) symbol language activated

Prior to the system's implementation, the various recipe steps were analyzed and categorized according to their interaction style. In a typical recipe, the student would first gather the ingredients and utensils necessary, then mix them together or with kitchen appliances, or just interact with the kitchen appliances. A total of 32 recipes provided by the RC educators, were studied and analyzed. Based on this dataset, the steps of a recipe have been classified as follows:

- card retrieval (e.g., "Take the cutting board and the knife"), where the student is expected to pick a card depicting a recipe material or utensil and place it on the table
- move of an element to another slot (e.g., "Put the colander in the sink"), where a virtual element should be selected through the pointing device and placed in an interactive slot, by pointing device selection
- the combination of elements (e.g., "Put the milk in the cup"), during which two virtual elements are expected to be combined by first selecting the one and then the other via the pointing device
- setting of a specific value to an appliance (e.g., "Set the timer to 10"), where the student selects the appliance and interacts with the UI that appears
- plug or unplug a device (e.g., "Plug the toaster"), by selecting it
- turn a device on and off (e.g., "Turn on the toaster"), by selecting it, and
- wait for a process to end (e.g., "Wait for the water to boil"), where no interaction is expected.

In some cases, a step may have a repetitive nature (e.g., "Put 5 glasses of water in the pot", where the student is expected to select the faucet, the glass and the pot, five times). This kind of step can be deducted into five "combination" type steps, while repetition is assisted by a bar with the appropriate number of slots, being filled with color as the student advances with the repetition.

At the beginning of a recipe's step, a text-to-speech service announces the related instruction. If the step asks for the retrieval of some ingredients or utensils, a placeholder element appears on the table signifying that cards placed there will be recognized by the system. The student must find and place the correct cards on the placeholder (Fig. 4(a)) and then confirm the answer by selecting the arrow button at the bottom right corner. Before continuing to the next step, each card's virtual counter-part appears on the system's virtual table, which can be manipulated with the pointing device. An example

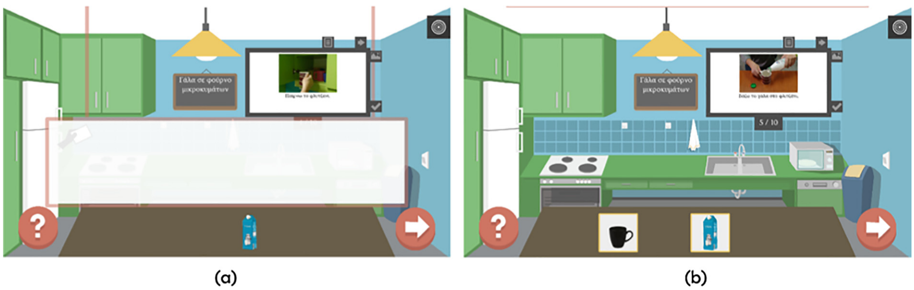


Fig. 4. The *Let's Cook* game (a) "Take the cup" step and (b) "Put milk in the cup" step

shown in Fig. 4(b) depicts a combination step, the "Put milk in the cup", in which the student has selected the milk and the cup, and is ready to confirm the answer.

Context-aware help is available to the student in each step, by tapping on the question-mark button at the bottom left corner. If help is asked, the elements to be selected are highlighted, indicating the slots that the student must select. When the student indicates that they are ready, the answer is assessed by the system and appropriate feedback is provided. In case of wrong answer, the system provides auditory advice about the step and encourages the student to try again. On the other hand, if the step was completed successfully, a clapping sound rewards the student and a video of an actual execution of this step in the real kitchen of the RC is played in a pop up window. Once the recipe is completed, the entire video is played, while a review of all the steps as they were executed by the student is readily available for the teacher.

An important feature of *Let's Cook*, is the adaptability it supports based on the profiling mechanisms implemented for the *CocinAR* system. More specifically, students are registered by their teachers, creating a simple user profile, with the child's name and functionality level, as it has been assessed by the RC educators. Once a user is logged in, additional settings can be defined referring to the content representation modalities that should be employed (text, images, speech, symbols), font size, audio volume, and sound effects activation. Speech instructions are provided using text-to-speech, which is also customizable and supports three different speech speed rates, allowing volume adaptation and complete deactivation. The most recent settings used by a particular student are retrieved and applied for every new interaction, thus personalizing the system according to the students' current needs. Furthermore, through the profiling mechanism, the teacher can have access to game history information, facilitating recognition of each student's strengths and weaknesses and planning of interventions. Game history information includes the number of times a game was played, the number and description of all the wrong answers in each game session, the number of times a student has asked for help, as well as the score achieved.

Finally, one of the most important aspects of the game is that it supports configurable interface and dynamic content retrieval, making its functionality completely context free and reusable. In more details, all word phrases and multimedia are externally defined so they can be easily replaced, while the graphics interface is customizable. All recipe related content is dynamically retrieved from a design studio, the *ConstrAct* editor [18]. This provided configurability allows the system to be repurposed for exercises with different content, preserving its functionality and interaction methods.

5 Evaluation

The evaluation of the *Let's Cook* game was conducted at the RC kitchen area, where children are taught to prepare simple meals, with the goal to identify any interaction problems, and to draw general conclusions about the system and the user experience.

Twelve children (Table 1) and their educator participated in the evaluation. Children were selected to participate in collaboration with the educator, so as to cover an as-wide-as-possible range of developmental disabilities. Their functionality level varied from 3-5 in a 1-5 scale, with 5 being the most functional. Regarding the demographic data of the participants, the user sample consisted of six boys and six girls, aged from 9 to 17 years old. Because of their age, their guardians were requested to sign a consensus form, indicating their consent regarding their child's participation.

Table 1. Participants of the evaluation

Disability	Children	Functionality level	Reading skills
Autism spectrum disorders	2	4, 5	Yes, Yes
Down syndrome	3	3, 4, 5	Somewhat, Somewhat, Yes
Psychomotor impairment	2	4, 4	Yes, No
Mental retardation	2	4, 5	No, No
Cerebral palsy	3	3, 4, 4	Somewhat, Somewhat, Yes

The evaluation method used was that of user observation. For this purpose, two usability experts were present during each session and were situated at a distance from the child and the teacher to avoid distracting the child. During each session, the experts paid close attention to the way both the child and the teacher interacted with the system. The observation sessions were spread along a five days' interval, keeping each session's duration the same as in their usual class. At the end of each session, the teacher was asked clarifying questions by the evaluator, about specific actions or errors of the current participant. At the end of all the sessions, the educator was queried with more general questions, concerning the students' overall interaction with the system, their likes and dislikes and the system's value for the teachers, following the semi-structured interview method [19]. The observation results and the answers collected during the experiment were recorded, kept and analyzed following an anonymity protocol by the same two evaluators, who were responsible for the evaluation sessions.

Analysis of the results[1] highlighted that the participants had no problem comprehending the layout of the game and the overall style of interaction. They knew where to expect the various elements to appear and what they had to do in order to check their answer and move to the next step. Additionally, the graphic elements were appropriate for all the participating children and they could clearly understand whether an item was selected or not. Regarding the input multimodality, it was observed that all children could switch between the two interaction styles with remarkable ease. They comprehended immediately the "fetch" paradigm of the card retrieval steps and could recognize the ones that required interaction via the pointing device. The multimedia output was verified as a necessary part of the system, facilitating a personalized content provision, able to meet the students' diverse needs and dexterities. Moreover, the multimedia output allowed the practice of additional, apart from cooking, skills, such as reading and listening comprehension, granting a multi-purpose character to the system.

[1] The detailed evaluation questions and their statistical analysis will be reported elsewhere due to space limitations.

However, a difficulty in the comprehension of long sentences was observed, leading to the suggestion that the system should be able to present long instructions gradually, pausing after each requested item until it is successfully selected to reveal the next part of the sentence. Another issue related to instructions' presentation was observed in the steps where the medium necessary for the step was implied rather than clearly stated (e.g., "Put cereal in the bowl", where the spoon to be used as a medium was omitted), so the students were likely not to select it. To address this issue, the system could help the student by asking an indicative question, currently asked by the teacher.

The fundamental question of whether the children actually liked the system was answered both by the teacher and also by the observed behavior of the children. According to the educators, the system is appealing to their students, keeping them focused to the task at hand, while at the same time it facilitates the better cognitive comprehension of the recipe, bringing students closer to their educational goal, i.e., its complete memorization.

6 Conclusions and Future Work

This paper has presented *Let's Cook*, an educational game which aims to get children with cognitive impairments acquainted with elementary meal preparation notions, while employing novel input methods through cards and pointing devices on a tabletop surface. The system supports multimedia output, personalized according to each individual child's needs and preferences. The usability evaluation conducted after the final installation of the system, which involved children with cognitive impairments, indicated that the interaction with *Let's Cook* enriches the user experience, making it more engaging for children and thus facilitates the educational purpose of their teachers. Since this technology is completely new to both sets of users, questions were raised in concern to whether the teachers would find it conducive and supportive to their learning model. To address this question future work will involve a longitudinal study, involving participants over a longer period of time, in order to draw more concrete conclusions about the educational value of the system.

Acknowledgments. This work is supported by the FORTH-ICS internal RTD Programme 'Ambient Intelligence Environments'.

References

1. Braddock, D., Rizzolo, M.C., Thompson, M., Bell, R.: Emerging technologies and cognitive disability. J. Spec. Educ. Technol. **19**(4), 49–56 (2004)
2. Scherer, M.J.: Assistive Technologies and Other Supports for People with Brain Impairment. Springer, New York (2011)
3. Salah, A.A., Schouten, B.A., Göbel, S., Arnrich, B.: Playful interactions and serious games. J. Ambient Intell. Smart Env. **6**(3), 259–262 (2014)

4. Göbel, S., Hardy, S., Wendel, V., Mehm, F., Steinmetz, R.: Serious games for health: personalized exergames. In: Proceedings of the 18th ACM International Conference on Multimedia, pp. 1663–1666. ACM (2010)
5. Lindley, C.A., Sennersten, C.C.: Game play schemas: from player analysis to adaptive game mechanics. In: Proceedings of the 2006 International Conference on Game Research and Development, pp. 47–53. Murdoch University (2006)
6. Dunleavy, M., Dede, C., Mitchell, R.: Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. J. Sci. Educ. Technol. **18**(1), 7–22 (2009)
7. Ntoa, S., Leonidis, A., Korozi, M., Papadaki, E., Adami, I., Margetis, G., Antona, M., Stephanidis, C.: Analysis and design of three multimodal interactive systems to support the everyday needs of children with cognitive impairments. In: Antona, M., Stephanidis, C. (eds.) UAHCI 2015. LNCS, vol. 9177, pp. 637–648. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20684-4_61
8. Mechling, L.C.: Assistive technology as a self-management tool for prompting students with intellectual disabilities to initiate and complete daily tasks: a literature review. Educ. Train. Dev. Disabil. 252–269 (2007)
9. Lopresti, E.F., Mihailidis, A., Kirsch, N.: Assistive technology for cognitive rehabilitation: state of the art. Neuropsychol. Rehabil. **14**(1–2), 5–39 (2004)
10. Lopresti, E.F., Bodine, C., Lewis, C.: Assistive technology for cognition [Understanding the Needs of Persons with Disabilities]. Eng. Med. Biol. Mag. IEEE **27**(2), 29–39 (2008)
11. Brandão, J., Cunha, P., Vasconcelos, J., Carvalho, V., Soares, F.: An augmented reality gamebook for children with autism spectrum disorders. In: The International Conference Focused on e-Learning in the WorNplace (2015)
12. Silva, G.F.M., Raposo, A., Suplino, M.: PAR: a collaborative game for multitouch tabletop to support social interaction of users with autism. Procedia Comput. Sci. **27**, 84–93 (2014)
13. Chang, Y.J., Kang, Y.S., Huang, P.C.: An augmented reality (AR)-based vocational task prompting system for people with cognitive impairments. Res. Dev. Disabil. **34**(10), 3049–3056 (2013)
14. Mansor, E.I., De Angeli, A., De Bruijn, O.: Little fingers on the tabletop: a usability evaluation in the kindergarten. In: 2008 3rd IEEE International Workshop Horizontal Interactive Human Computer Systems, TABLETOP 2008, pp. 93–96. IEEE (2008)
15. Marco, J., Cerezo, E., Baldassarri, S., Mazzone, E., Read, J.C.: Hybrid games: designing tangible interfaces for very young children and children with special needs. In: Nijholt, A. (ed.) More Playful User Interfaces, pp. 17–48. Springer, Singapore (2015). https://doi.org/10.1007/978-981-287-546-4_2
16. Jang, H.B., Kim, J.W., Lee, C.W.: Augmented reality cooking system using tabletop display interface. In: International Symposium on Ubiquitous VR, p. 1 (2007)
17. Zidianakis, E., Antona, M., Paparoulis, G., Stephanidis, C.: An augmented interactive table supporting preschool children development through playing. Proc. AHFE Int. **2012**, 21–25 (2012)
18. Poutouris, E., Korozi, M., Leonidis, A., Louloudakis, N., Stephanidis, C.: ConstrAct: an educator-oriented design studio for ambient educational games. In: Proceedings of the 9th Annual International Conference on Education and New Learning Technologies, pp. 8475–8484, EDULEARN 2017 Proceedings (2017)
19. Drever, E.: Using Semi-Structured Interviews in Small-Scale Research. A Teacher's Guide (1995)
20. Lancioni, G.E., O'Reilly, M.F., Seedhouse, P., Furniss, F., Cunha, B.: Promoting independent task performance by persons with severe developmental disabilities through a new computer-aided system. Behav. Modif. **24**(5), 700–718 (2000)

21. Mechling, L.C., Gast, D.L., Seid, N.H.: Using a personal digital assistant to increase independent task completion by students with autism spectrum disorder. J. Autism Dev. Disord. **39**(10), 1420–1434 (2009)
22. Miyawaki, K., Sano, M., Yonemura, S., Matsuoka, M.: A cooking support system for people with higher brain dysfunction. In: Proceedings of the ACM Multimedia 2009 Workshop on Multimedia for Cooking and Eating Activities, pp. 47–52. ACM (2009)
23. Ju, W., Hurwitz, R., Judd, T., Lee, B.: CounterActive: an interactive cookbook for the kitchen counter. In: CHI 2001 Extended Abstracts on Human Factors in Computing Systems, pp. 269–270. ACM (2001)

GHio-Ca: An Android Application for Automatic Image Classification

Davide Polonio, Federico Tavella, Marco Zanella, and Armir Bujari(✉)

Department of Mathematics, University of Padua, Padua, Italy
{davide.polonio,federico.tavella,marco.zanella.8}@studenti.unipd.it,
abujari@math.unipd.it

Abstract. Online social networks (OSN) have revolutionized many aspects of our daily lives and have become the predominant platform where content is consumed and produced. This trend coupled with recent advances in the field of Artificial Intelligence (AI) have paved the way to many interesting features, enriching user experience in these social platforms. Photo sharing and tagging is an important activity contributing to the social media data ecosystem. These data once labeled constitute a fruitful input for the system which is exploited to better the services of interest to the user. However, these labeling activity is imperfect and user subjective, hence prone to errors inherent to the process. In this paper, we present the design and the analysis of an Android app (namely GHio-Ca), an automatic photo tagging service relying on state-of-the-art image recognition APIs. The application is presented to the user as a camera app used to share pictures on social networks while relying on external services to automatically retrieve tags best representing the picture theme. Along with the system description we present a user evaluation involving 30 subjects.

Keywords: Online social networks · Social media sensing
Computer vision · Android · Image recognition

1 Introduction

Social media has become a ubiquitous part of everyday life. The amount of data being published through these services contains valuable potential information which can be exploited by algorithms and put to good use in order to provide new and innovative services to the users [1–3]. However, extracting the semantics from the data is generally a hard problem and user provided metadata could aid to better contextualize and infer information.

One such category of data are photos published on social networks which are generally associated with a description and some hashtags labeling them. The latter could be personal words that a person associates to a certain photo (e.g. feelings, person names, places) but they could also be used in order to describe the content of the photo. All these pieces of information can be really useful

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 248–257, 2018.
https://doi.org/10.1007/978-3-319-76111-4_25

to train algorithms and to create datasets used in order to recognize images: in fact, these data, which are freely available, are posted by people that manually label and describe a specific photo.

However, user provided hashtags are often imprecise, subjective [4]. As an alternative one could rely on image recognition services to automatically tag photos prior to sharing on social network sites. This has the potential benefit of employing unbiased metadata, providing a more useful feedback to the services that rely on them. In specific, image recognition services process an image and return a set of labels associated to that picture. Results vary in precision depending on the photo quality, subject and also on the algorithm being employed.

In this context, we propose *GHio-Ca* (*Giving Hashtags In Order to Classify Automatically*), an Android application allowing people to take photos or choose pictures which are automatically processed by some image recognition service. Our application was designed with quality of service and usability requirements in mind and its motivation is twofold: (i) automatically and transparently provide useful and meaningful information aiding the user and (ii) create an unbiased image dataset used to train image recognition algorithms. In specifics, in order to achieve our purpose, we rely on the following image recognition services: (a) Computer Vision API by Microsoft Azure [5], (b) Visual Recognition by IBM Watson [6], (c) Google Reverse Image Search [7] and (e) Imagga [8]. To asses our proposal, we undertook a user study evaluating the different services in terms of result accuracy and user satisfaction.

This paper is organized as follows: Sect. 2 provides some background information on the field of artificial intelligence and image processing along with an overview of the current state of the art of computer vision applications. Section 3 describes the application design and implementation, while Sect. 4 provides a comparison analysis of the different image recognition APIs we exploited. Finally, in Sect. 5 conclusions are drawn.

2 Background and Related Work

Machine learning is the field of artificial intelligence responsible for learning from data without being explicitly programmed to do so. In this way, the results provided by the algorithm do not depend on how data is processed, but rather on the data itself. Broadly there are two methods for training an algorithm on a dataset: *supervised* and *unsupervised* learning. In supervised learning, each data in the dataset is associated to a label. Consequently, we are able to train the algorithm based on examples and the objective is to predict the label for future data. In unsupervised learning, there is no such association. Thus, to determinate if two examples are referring to the same result, one relies on a *similarity measure* (e.g. if we represent data using vectors, one possible measure of similarity is the cross product).

Pattern Matching is the branch of machine learning responsible for detecting patterns and regularities in data. Typically, it is implemented through supervised learning approach: the dataset is composed by examples with an associated label. In this way, the algorithm learns which are the features of a specific pattern. One of the most used method in machine learning for image classification is *Deep learning*: this technique uses an Artificial Neural Network (*ANN*) - which is a Neural Network (*NN*) composed by more than one hidden layer, as shown in Fig. 1 - in order to classify an image, based on similarity measures (unsupervised) or training examples (supervised). Typically, NNs use a *backpropagation* algorithm, composed of two phases: *feed forward* and *backward propagation*. First, the image is decomposed in a vector-like representation and at a second stage, during feed forward phase, the vector is fed as input to the NN and it is computed. Finally, during back propagation, the result of feed forward phase is compared to the real label; in case of mismatch, the wrong parts are backpropagated into the NN in order to compensate the wrong implementation. This process is repeated for each entry in the dataset.

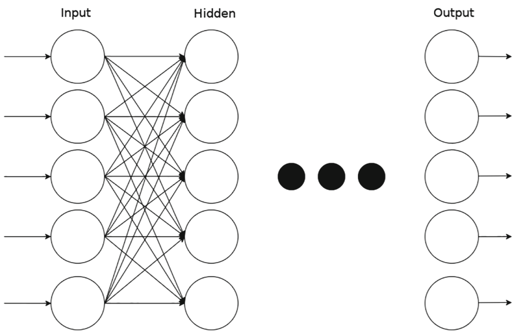


Fig. 1. Architecture of a Neural Network

Computer Vision is a branch of machine learning which is rapidly increasing, finding a fertile ground in many applications. One of the most obvious is *image processing*, which involves applying changes to an image using an algorithm. In [9], researchers developed a DNN used to remove rain drops from images. As in this case, many other changes can be applied to images: increasing size with minimum quality loss, removing objects, improving image resolution, and so on. However, computer vision contains also another kind of application: *image detection*. Using NNs, algorithms are able to detect pedestrians in images [11] (which is fundamental for self-driving cars) and also detect anomalous behaviors in crowds [10]. Being able to detect shapes (e.g. a person) and irregular patterns can also help in terrorism detection and prevention.

On the other hand, image classification has two main problems: *time* and *cost*. This task can be performed by a machine but, in order to do so, the algorithm needs to be previously trained with a supervised learning approach. Consequently, someone has to label different images. This task is typically done by people who are paid to execute it. Differently from machines, humans are far from fast to perform image labeling; consequently, in order to obtain a large dataset, a lot of time is required. As in other cases in Computer Science, *parallelization* can improve the performance of this job: if a person needs two seconds to label an image, one thousand people can provide one thousand labeled images in the same amount of time. One good example of this principle is the database ImageNet [13], “a large-scale ontology of images built upon the backbone of the WordNet structure [14]. This database can be used as benchmark and improvement tool for computer vision algorithms. However, in our opinion even this last case lacks of two fundamental properties: *usability* and *zero-day learning*. Obviously, accessing to such database in order to retrieve information and perform pattern recognition is not a task which can be performed by a common user (i.e. a person who does not have any skill in computer science and programming) without a proper user interface. Furthermore, this database needs to be populated in the first place: this task requires paying people to do so.

In this context, our application aims to tackle these issues, providing a usable interface for image recognition by exploiting smartphones to build a database from third-party services. At the same time, the application is useful to end-users, providing an automatic hashtag feature easily integrated with social network platforms.

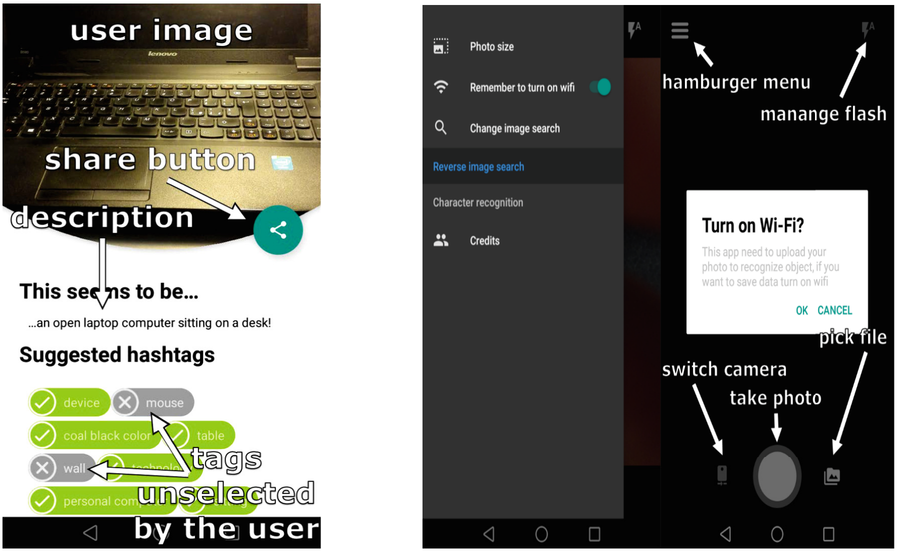
3 GHio-Ca

In this section we provide a description of some salient features of our application, mainly concerning architectural and implementation choices made during the development process.

3.1 Architectural Design

GHio-Ca is an Android application that provides a user the possibility to automatically label photos and share them through social network sites. The application embeds some camera features which allow the capture of photos which are successively uploaded in order to be processed by an image recognition engine. Also, it allows the user to pick a photo from the local storage and start the classification process on it. Regarding the API version compliance, we imposed API level 21 (Android Lollipop) as a minimum requirement. The source code of the project can be found at [15].

In Fig. 2 are shown the main views present in the app ranging from the camera acquisition, configuration and results view. From an architectural design viewpoint, GHio-Ca is composed by three loosely coupled modules (Fig. 3): (i) the first one manages network connectivity, uploads the photo to the server and makes requests to different services in order to make image recognition or the translation of certain pieces of text; (ii) the camera module, that manages the photo capturing process and storage on the devices, and (iii) the module that glues all the pieces together. These modules are maintained as loose coupled as possible, in order to make it easier to change the used services without making important changes to the overall application (e.g. changing the process of capturing a photo without modifying the networking module).



(a) `ResultActivity` for image recognition

(b) `CameraPreviewActivity`, with menu open and close

Fig. 2. Screenshot taken from the application

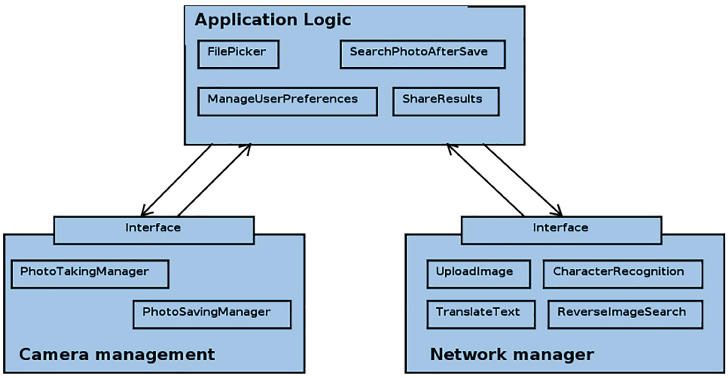


Fig. 3. GHio-Ca architecture

All communication with external services is done through a background worker thread listening from requests coming from the UI part. Whenever a communication error with an external service occurs, the user is presented with an error notifying the undesirable outcome. When no errors occur the results are skimmed based on the probability that a word (or bag of words) could be correlated to the user picture. Depending on the capabilities offered by the external service the application relies on, one could set a correlation probability threshold. This feature gives the possibility to present the user more valid results from which to chose from. This value is configurable through the application settings menu and is set to 70%. On the other side, if the remote service does not provide this capability, the user is presented with the entire bag of words (hashtags) from which to chose from.

The duty of the network module goes beyond the sole responsibility of handling the interaction with external services. Due to limitations of some services and to reduce network usage, we made the chose to previously upload the photos to a server, and send the URL of the image to the different services we rely on. With this approach, we reduce bandwidth usage: the user needs to upload the photo only once and can take advantage of the URL to make the recognition with more services.

The camera module, instead, manages the photo acquisition and persistence process. These activities are fulfilled using the CameraFragment library. We employ fragments in order to show the user a preview of the photo and, when the user taps on the button to take the picture, saves it in a specific folder. Successively, once the photo is acquired, the recognition process can start. This module also gives the user the possibility to use either the frontal of back camera view, manage flash (turn on/off or choose automatic option) and enable the user to choose the size of the photos.

The last module interconnects the prior modules by employing custom Android components for inter-module communication. Moreover, it manages the application sharing process: as a matter of fact, after the recognition process finishes, the application gives the user the possibility to share the photo on different social networks sites or via other means (e.g., email). For some of them (e.g., Facebook, Twitter, Instagram, Whatsapp, Linkedin) there is a specific implementation, while for the others, the default Android support is used.

3.2 Recognition Services

Libris (Library for Reverse Image Search) is a library that we developed in order to simplify the image recognition process. The library aims to make easier calling the different services utilized, in order to fulfill the image or character recognition, defining an interface for the results.

All the requests return a response in `JSON` format, which is parsed. The more relevant fields are embedded in an object, different for every service, which is returned as result. If some error occurs during the request an `Exception` is thrown.

For image recognition Libris gives the possibility to use all the services listed in Sect. 1. In order to provide the Google Reverse Search Image service, we programmatically search on Google the image, choosing the best returned result.

Before delving into the evaluation part, we discuss some limitations and problems encountered with the adopted recognition services. The Azure Image Recognition is the core of GHio-CA and it works pretty well. Unfortunately, we had some problems with the service provider: firstly, it banned our first account; secondly it did not accept our student subscription that would have enabled us to take advantage of it. Finally, without any notification, Microsoft restored our first account and we were able to use the Image Recognition service.

We also adopted an Image Recognition service from IBM to obtain more tags from a shot performed by the user. Although we did not have problem with the service per-se, we found a bug in the Watson SDK related to the data format used, that we where able to fix.

4 Results

Through this section we discuss the evaluation strategy and the outcome of the field trial. The main purpose of the trial was to asses the validity of each recognition service by essentially computing some statistics on the returned bag of words. To this end, we distributed GHio-Ca to 30 subjects and asked to deselect tags that did not concern shots they had taken. Finally, they had to send the result to us.

Our approach was to make the application usage as simple as possible for the subject to contribute. Prior to the trial we prepared (i) a video tutorial in which we showed what they had to do, (ii) wrote a mini-wiki where the experiment was explained, (iii) modified the application accordingly. In line with our recommendations, each subject was asked to take 5 photos. The final number of photos that we received is 150.

For the evaluation part, we distributed to the subjects a modified version of the application in order to contain and limit the API requests made to each individual service. Indeed, there is an upper limit to the number of requests imposed by each service. Also, instead of sharing the classification outcome through a social media platform, each result is sent to a pre-configured email address.

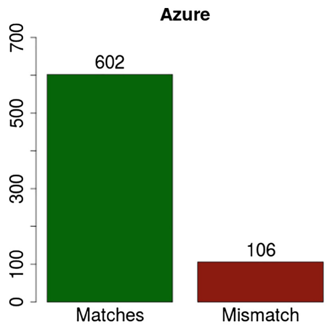
4.1 Graphical Analysis

At the end of the testing phase, we collected all the e-mails received, we counted them and we built several graphical representations that we explain as follows.

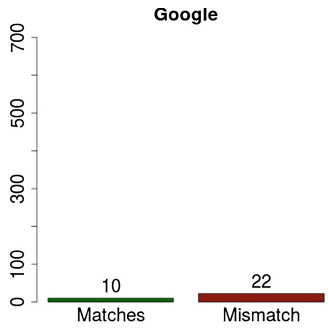
In first place, Azure was the service that gave us the best results with 85% of correct tags (Fig. 4a), on the other hand Google Reverse Image Search (Fig. 4b) was the worst one (with only 31% of matches), but we have to advocate that it was not designed to provide tags from/to user images, thus we did not expect good results from this service.

Imagga (Fig. 4c) and Watson (Fig. 4d) (respectively a service we found online and an IBM Image Recognition service) didn't stand out as we expected: for the first one we had to filter a good chunk of tags since a lot of them had a low confidence of correlation with images and the second one provided a lot of tags that turned out to be wrong. The percentages of good tags are 78% for Imagga and 65% for IBM Watson.

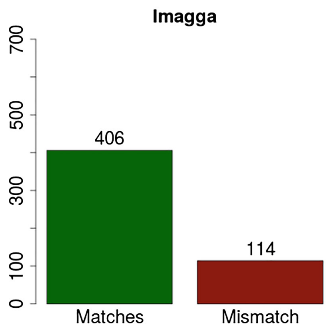
In Fig. 5 we can see the overall representation, with a percentage of 75% of correct tags, that we consider to be a good result.



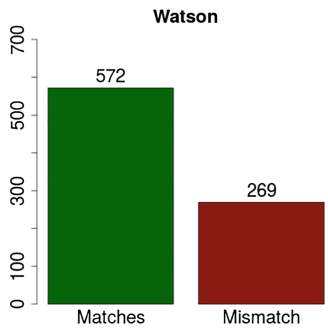
(a) Azure tagging results (708 tags provided)



(b) Google tagging results (32 tags provided)



(c) Imagga tagging results (520 tags provided)



(d) Watson tagging results (841 tags provided)

Fig. 4. Frequencies histogram of matches/mismatches

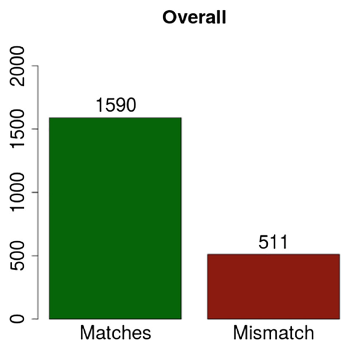
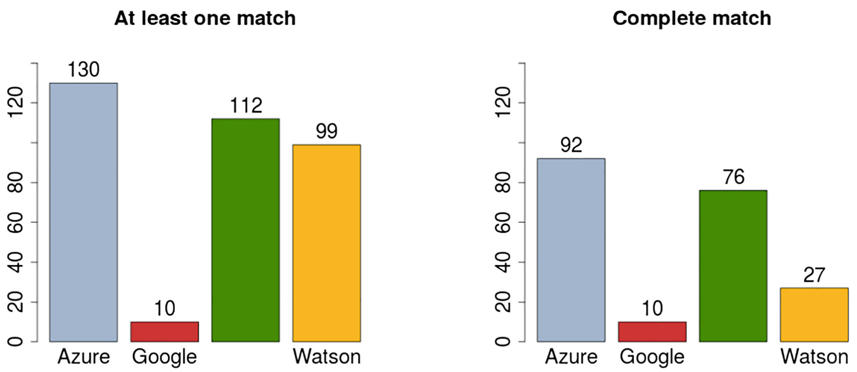


Fig. 5. Histogram of all tagging results (2101 tags provided)



(a) Histogram of all searches that give back at least one match

(b) Histogram of all searches that give back only correct tags

Fig. 6. Image classification details

In Fig. 6a we can see the number of times that services gave back at least one correct tag. Even in this case, Azure was the service that had the best performance but also Watson and Imagga gave back results that are good enough.

Finally, we present in Fig. 6b the number of times that a service returned a bunch of tags that were all confirmed as matches by a user. In this case Azure and Imagga had the best results. In this case even Google Reverse Image Search has a positive result, because it gives to the user only one tag that could be accepted or not.

5 Conclusion

In this article, we presented GHio-Ca, an mobile application enabling users to access to different image recognition APIs. The application is used for automatic image labeling (i.e., hashtagging) and can be easily integrated to the social network ecosystem. GHio-Ca has a higher purpose that of speeding up the creation of training/testing datasets for machine learning algorithms in the context of the so called *crowd learning*. Our solution could be a step towards overcoming the issue present in image recognition libraries (Sect. 2) given that everyone can use an application on its smartphone (no usability problems) and it is possible to get rid of external image recognition services after getting over a critical threshold (solving the *zero-day learning* problem). We tested different recognition APIs, inferring that the best one is Computer Vision API by Microsoft Azure [5], followed by Imagga [8] and Watson [6]. As discussed in Sect. 4, Google Reverse Image Search [7] is tailored to label images, hence the justification of the low accuracy.

References

1. Bujari, A., Furini, M., Laina, N.: On using cashtags to predict companies stock trends. In: Proceedings of IEEE CCNC, Las Vegas, NV, USA, pp. 25–28 (2017)
2. Roccetti, M., Salomoni, P., Prandi, C., Marfia, G., Mirri, S.: On the interpretation of the effects of the infliximab treatment on Crohn's disease patients from Facebook posts: a human vs. machine comparison. Netw. Model. Anal. Health Inf. Bioinf. **6**(1), 11 (2017)
3. Bujari, A., Licar, B., Palazzi, C.E.: Road crossing recognition through smartphone's accelerometer. In: Proceedings of IFIP WD, Niagara Falls, ON, pp. 1–3 (2011)
4. Bujari, A., Ciman, M., Gaggi, O., Palazzi, C.E.: Using gamification to discover cultural heritage locations from geo-tagged photos. Pers. Ubiquit. Comput. **21**(2), 235–252 (2017)
5. Computer vision API by Microsoft Azure. https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/
6. Visual recognition by IBM Watson. https://www.ibm.com/watson/developercloud/visual-recognition.html
7. Google reverse image search. https://images.google.com/
8. Imagga. https://imagga.com/
9. Fu, X., Huang, J., Ding, X., Liao, Y., Paisley, J.: Clearing the skies: a deep network architecture for single-image rain removal. IEEE Trans. Image Process. **26**(6), 2944–2956 (2017)
10. Sabokrou, M., Fayyaz, M., Fathy, M., Klette, R.: Deep-cascade: cascading 3D deep neural networks for fast anomaly detection and localization in crowded scenes. IEEE Trans. Image Process. **26**(4), 1992–2004 (2017)
11. Cao, J., Pang, Y., Li, X.: Learning multilayer channel features for pedestrian detection. IEEE Trans. Image Process. **26**(7), 3210–3220 (2017)
12. Krishna, R., et al.: Visual genome: connecting language and vision using crowdsourced dense image annotations. Int. J. Comput. Vis. **123**(1), 32–73 (2017)
13. Russakovsky, O., et al.: ImageNet large scale visual recognition challenge. Int. J. Comput. Vis. **115**(3), 211–252 (2015)
14. Deng, J., Dong, W., Socher, R., Li, F.: ImageNet: a large-scale hierarchical image database. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 2009
15. GHio-Ca source code repository. https://github.com/Augugrumi/ghioca
16. Mason, W., Suri, S.: Conducting behavioral research on Amazon's Mechanical Turk. Behav. Res. Methods **4**(1), 1–23 (2012)
17. LeCun, Y., et al.: Backpropagation applied to handwritten zip code recognition. Neural Comput. **1**(4), 541–551 (1989)
18. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. In: Proceedings of IEEE, vol. 86, no. 11, pp. 2278–2324, November 1998

The Analysis of Influential Users Evolution in Microblogging Social Networks (Extended Abstract)

Giambattista Amati[1], Simone Angelini[1], Giorgio Gambosi[2], Gianluca Rossi[2], and Paola Vocca[3](✉)

[1] Fondazione Ugo Bordoni, Rome, Italy
{gba,sangelini}@fub.it
[2] University of Rome "Tor Vergata", Rome, Italy
{giorgio.gambosi,gianluca.rossi}@uniroma2.it
[3] University of Tuscia, Viterbo, Italy
vocca@unitus.it

Abstract. In this paper, we study the evolution of the most influential users in the microblogging social network platform Twitter. To this aim, we consider the *Dynamic Retweet Graph (DRG)* proposed in [3] and partially analyzed in [2,4]. The model of the evolution of the Twitter social network is based on the retweet relationship. In a DRGs, once a tweet has been retweeted the last time all the edges representing this tweet are deleted, to model the decay of tweet life in the social platform.

We consider the following measures of centrality: *degree, closeness, and pagerank-centrality* which have been widely studied in the static case. Here we analyze them on the sequence of DRG temporal graphs with special regard to the distribution of the 75% most central nodes.

We derive the following results: (a) in all cases the closeness measure produces many nodes with high centrality, so it is useless to detect influential users; (b) for the other measures almost all nodes have null or very low centrality and (c) the number of vertices with significant centrality are often the same; (d) the above observations hold also for the whole DRG and, (e) central nodes in the sequence of DRG temporal graphs have high centrality in static graphs.

Keywords: Graph analysis · Social media · Twitter graph
Retweet graph · Graph dynamics · Centrality

1 Introduction

One of the fundamental and most studied features in a social network is the detection of central nodes, which can usually be considered as the *most important* nodes [6,7,12]. Centrality is widely-used for measuring the relative importance

This work was conducted in the Laboratory of Big Data of ISCOM-MISE (Institute of communication of the Italian Ministry for Economic Development).

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 258–268, 2018.
https://doi.org/10.1007/978-3-319-76111-4_26

of nodes within a graph and has many applications: in social networks to determine the most influential or well-connected people; in the Web graph to rank pages in a search; in a terrorist network, to detect agents that are critical for facilitating the transmission of information; for the dissemination of information in P2P Networks, Decentralized Online Social Networks and Friend-to-Friend Network [10].

There is a plethora of centrality definitions: degree centrality [16], closeness centrality [5], graph centrality [13], stress centrality [17], betweenness centrality [11], each one of them useful to detect specific properties and with significantly different computational costs. Here we consider four of them: the *degree*, *closeness*, *betweenness*, and *PageRank*-centrality.

Degree centrality, i.e. the degree d_v of a vertex v, is the simplest measure of centrality: it just takes into account how many direct, "one hop" connections each node has to other nodes of the network, hence it can be applied to detect popular individuals, agents who are likely to hold most information or individuals who can quickly connect with the wider network. The degree centrality is very cheap to compute but, being a purely local notion, it is often unable to recognize the relevance of certain nodes.

One of the most popular measures, but computational expensive for large graphs, is betweenness-centrality. It detects nodes which act as "bridges" between other nodes in a network. It does this by identifying all the shortest paths and then counting how many times each node falls on one. Betweenness centrality is suitable for finding vertices who influence flows (such as information flow) in the network.

A third measure considered below is closeness-centrality, which, after computing the set of all-pairs shortest paths, assigns each node a score based on the number of shortest paths to which it belongs. This definition of centrality is useful for quickly finding the agents who are in good position to influence the entire network but in a highly connected network often most nodes have a similar score.

Finally, Pagerank-centrality was introduced in [8] and it recursively quantifies a "value" or the PageRank of a node based on: (i) the number of links it receives, (ii) the link propensity of the linkers (that is, the number of outgoing links of each in-going node), and (iii) the centrality of the linkers, that is their PageRank.

In order to analyze the evolution the influential users we study the distribution of the centrality measures on a model of the Twitter network, the *Dynamic Retweet Graph (DRG)* proposed in [3] and partially analyzed in [2,4].

This model has two major features: (i) we consider the retweet graph since it allows to better represent relationships among users related to information flow in Twitter [14,15] and (ii) once a tweet has been retweeted for the last time all the edges representing that tweet are deleted, to model the decay of relevance of the tweet content.

The temporal model we consider coincides with the other ones in the growing phase. That is a new vertex is added whenever a new user starts or retweets a tweet, and a new directed edge (a, b) is inserted when a user a retweets for the first time a tweet of b, if an edge already exists then a timestamp is added to it. Conversely, the decreasing stage happens when a tweet is never retweeted again. Then, all vertices and edges not involved in other retweeting processes are deleted at once. As shown in previous experimentations [2,4], this evolutionary model better captures the information flow in Twitter.

DRGs seem to better represent the double nature of the Twitter platform: social network and news media [14,15].

For what concerns the use of the centrality measure to assess influential or authoritative users Kwak et al. [14] compared three measures of influence: in-degree centrality, PageRank centrality in the following/follower network and the number of retweets on Twitter. Cha et al. [9] compared three different measures of influence: in-degree centrality, the number of retweets and mentions on Twitter. The results indicate that users with high in-degree were not necessarily influential.

In this paper we study the evolution of the most influential users in the microblogging social network platform Twitter with respect to the above four centrality measures (betweenness, degree, closeness, and PageRank) and we analyze their behavior on the DRG evolutionary model of the retweet social networks proposed in [3].

We consider two different kind of data sets, first introduced in [1] and updated and refined in [3]: the *event driven* retweet graphs based on the events *Black Friday 2015* and the *World Series 2015* and the *Italian Sampling*, that is the *firehose* retweet graph, filtered by language (i.e. Italian) from the whole Twitter stream.

The four centrality measures are analyzed on three levels: (i) with respect to the sequence of DRG temporal graphs; (ii) with respect to the static cumulative graph, that is the graph that contains all nodes and edges and (iii) with respect to the kind of networks considered, that is *event driven* or the firehose.

We derive that the model proposed allows to detect the most authoritative users, since:

1. in all cases the closeness centrality provides too many central nodes, hence it is useless to detect influential users;
2. with regard the other measures, almost all nodes have null or very low centrality;
3. vertices with centrality values above 75% of the maximum is a small set and they are often repeated in the three centrality measures;
4. the above observations hold also for the static graphs (the whole DRG);
5. central nodes in the sequence of DRG temporal graphs have high centrality in static graphs.

2 DRG Temporal Graphs

In this paper we will use a definition of Dynamic Retweet Graph (DRG) slightly different from the one in [4].

A DRG graph $G = (V, E, \ell)$ is defined as follows: the set V of nodes are Twitter accounts and a directed edge $e \in E$ represents an interaction (a retweet) between two accounts. In particular, there is a directed edge from an account a to an account b, if a has retweeted at least one tweet of b, that can be itself already a retweet. Observe that user a may retweet more tweets of b. This edge information is implemented with a list $\ell(e)$ associated to every edge $e = (a, b)$ that contains pairs (i, t) where i is the id of a tweet and t is the timestamp in which a retweets i from b. The pairs of $\ell(e)$ are sorted by non-decreasing order of their timestamp.

From the data that we have collected in G we define, for all tweets i, the *date of death* of i (in short, $\mathtt{dod}(i)$) as the timestamp of the latest retweet of i. Formally,

$$\mathtt{dod}(i) = \max_{e \in E}\{t : (i, t) \in \ell(e)\}.$$

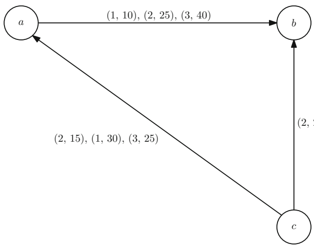
Consequently we define the *expiration date* of an edge e (in short, $\mathtt{ed}(e)$) as the time after which all tweets associated to e will be dead. Formally,

$$\mathtt{ed}(e) = \max\{\mathtt{dod}(i) : (i, t) \in \ell(e)\}.$$

On the contrary, the *creation date* of an edge $e = (a, b)$ (in short, $\mathtt{cd}(e)$) is the time when b retweets a for the first time, formally:

$$\mathtt{cd}(e) = \min\{t : (i, t) \in \ell(e)\}.$$

Let t be a timestamp, we define a *DRG temporal graph* at time t as the subgraph $G_t = (V_t, E_t)$ of the DRG G at time t defined as follows: E_t contains any edge e such that $\mathtt{cd}(e) \leq t \leq \mathtt{ed}(e)$; V_t is the set of nodes induced by E_t.



tweet id	$\mathtt{dod}(\cdot)$
1	30
2	25
3	40

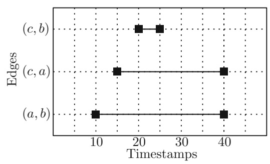


Fig. 1. On the left side, an example of a DRG retweet graph. Edges are labelled by pairs with the id of the tweet and the timestamp of the retweet. The center table shows the date of death of all tweets in the graph. On the right side, for each edge of G is represented its creation and expiration date.

For example if G is the retweet graph represented in the left part of Fig. 1, G_{30} contains edges (a, b) and (c, a) and the induced vertices since (c, b) expires at timestamp 25. For all $20 \leq t \leq 25$, G_t contains all edges of G.

3 Data Sets

For the experiments we use the same dataset as [3] that consists in two different classes of retweet graphs: the event driven retweet graph, filtered by topics about specific events (i.e. the Black Friday 2015 and the World Series 2015) and the sampling retweet graph, filtered by the Italian language from the whole Twitter stream. To obtain the Italian Twitter sample we use a list of the most used Italian stop words and the Twitter native selection function for languages. In Table 1 the size of the three graphs are shown. In Fig. 2 we show the evolution of the size of the three datasets over the period of observation. Note that the event-driven datasets (World Series and Black Friday) show a rapid growth close to the events, and then a slow decline. Differently, the Italian Sampling show a smooth and stable behavior, ignoring the border effects.

Table 1. Size of the dataset

	Black Friday	World Series	Italian Sampling
Vertices	$2.7e + 06$	$4.74e + 05$	$2.541739e + 06$
Edges	$3.8e + 06$	$8.40e + 05$	$1.3708317e + 07$
Tweets/edges	2.603	2.3	5.45
Tweets/vertices	3.66	4	29.4

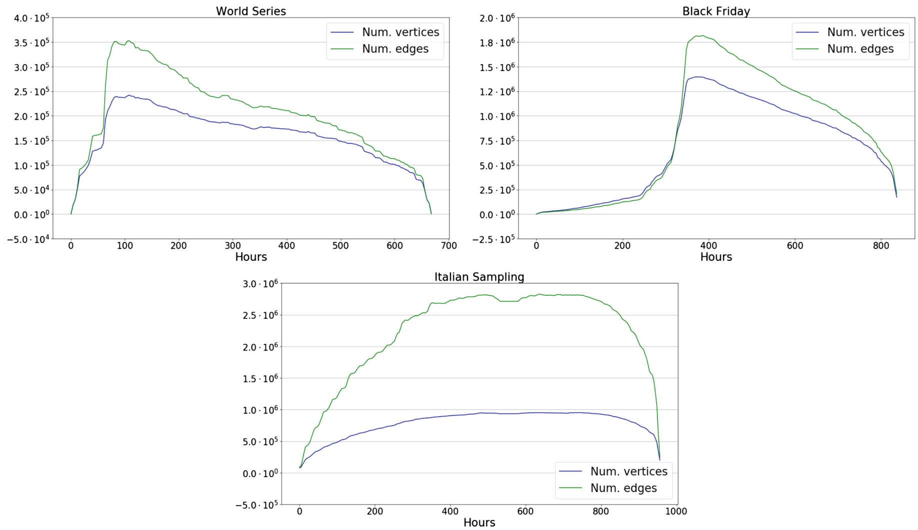


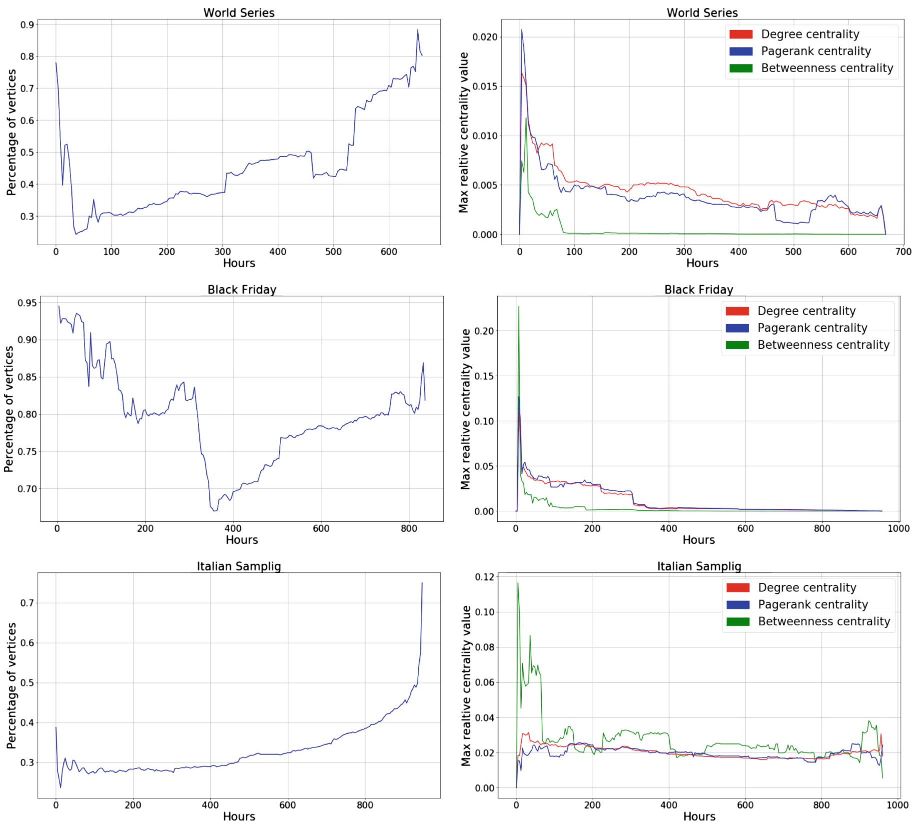
Fig. 2. Number of vertices (blue) and number of edges (green) of: World Series, Black Friday, and Italian Sampling, as functions of hours. (Color figure online)

4 Experimentation

For each graph G in our dataset, we consider the sequence of DRG temporal graphs $(G_{t_i})_{i \geq 0}$ where $t_{i+1} - t_i$ is 4 h. For each G_t we compute the four centrality values (betweenness, closeness, degree, and PageRank centrality) of each vertex of the graph.

Given the centrality measure c, the *relative centrality value* with respect to c of a vertex u is the ratio $c(u)$ and the maximum value of $c(\cdot)$.

Preliminary considerations. First of all, for each centrality measure $c(\cdot)$ and for each G_t, we consider the number of nodes with centrality values above 90% of the maximum. Figure 3(a) shows the behavior of the closeness centrality: observe that this value is almost always greater than 30%. This means that



(a) (b)

Fig. 3. (a) Trend over time of the ratio of nodes whose closeness centrality is above the 90% of the maximum. (b) The 99.9-th percentile evolution over time of the three relative centrality measures.

closeness centrality is not very suitable to determine the more influential nodes in the graph. Conversely, the other centrality measures (degree, betweenness, and PageRank) show an opposite behavior: excluding the first and last timestamp, 99.9% of vertices always have centrality values below 20% of the maximum. This is shown in Fig. 3(b) which shows the evolution over time of the three centrality values below which the 99.9% of all values fall (99.9-th percentile). Observe that, from Fig. 3(b) it results that the highest values are at the very beginning of time sequences, when there is still much instability. After that, values fall below 0.05.

Analysis of temporal graphs. From the previous observations it follows that if we restrict ourselves to the betweenness, degree and PageRank measures, the number of vertices for which the centrality value is meaningful is so small that we can study them one by one.

We say that a node is *central* (with respect to a centrality measure) if its centrality value is at least 75% of the maximum. Let G be a DRG, c be a centrality and t be a timestamp, we define $A_{G,c,t}$ as the set of central node of G_t with respect to c.

In Fig. 4 are shown the sets $A_{G,c,t}$ for the World Series (Italian Sampling and Black Friday are similar and are omitted for lack of space). In the y-axis are reported the vertex ids. Let us consider one of the diagrams in the figure relative

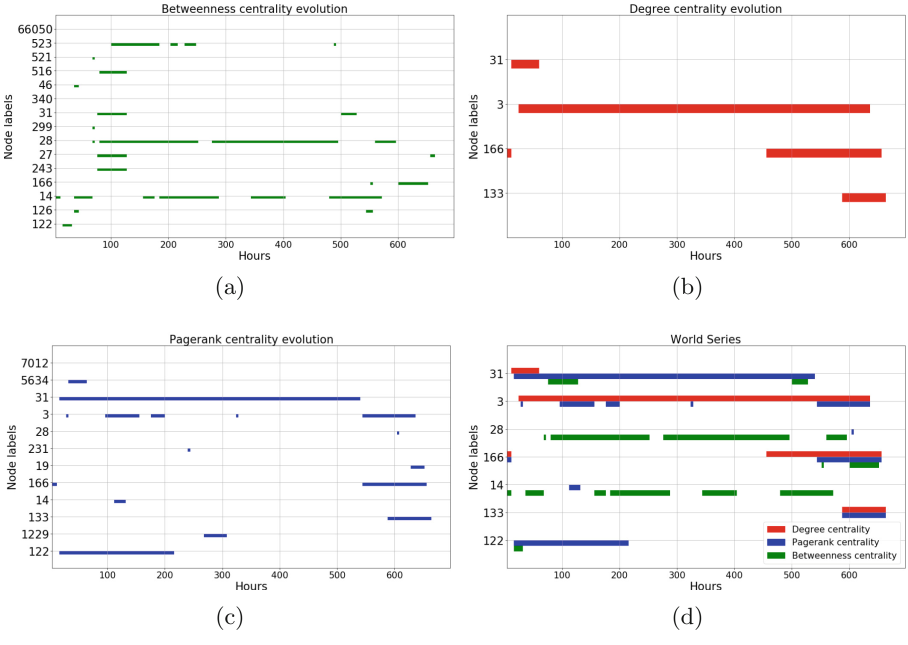


Fig. 4. Temporal evolution of $A_{G,c,t}$ for the world Series relative to centrality measures: betweenness (a), degree (b), and PageRank (c). Diagram (d) shows the overlapping of (a), (b), and (c) with respect at least two measures.

to measure c: a segment in correspondence to node u that intersects timestamp t means that $u \in A_{G,c,t}$. From the above analysis we get the following observations:

- For all datasets, the degree centrality always produces a total number of central nodes lower than the other measures. Conversely, betweenness centrality is the one that produces more.
- For all datasets and all the centrality measures, there are nodes that are central for a long time: this trend is more prominent for degree and pagerank centrality.
- Another important result that turns out is a significant overlap between the central vertices with respect to the three measures. In Fig. 4(d) the diagrams in Fig. 4(a)–(c) are merged by taking into account only vertices in common with at least two measures. For example vertex 31 is central for most of the time over the three measures.

Comparison with the static cumulative DRGs. The latest analysis involves the centrality measures of the static cumulative DRGs G representing the three datasets. Like DRGs temporal graphs, a large portion of vertices, varying form 28% (for World Series) to 50% (for Black Friday), have closeness centrality above 90% of the maximum, hence, we discard it.

On the contrary for the betweenness, degree, and PageRank centrality, almost all the nodes have centrality below 1% of the maximum. Table 2 shows, for each dataset and for each measure the percentage of vertices whose relative centrality value is at most 0.01.

Table 2. Percentage of vertices whose relative centrality value is at most 0.01.

World Series			Black Friday			Italian Sampling		
Between.	Degree	Pagerank	Between.	Degree	Pagerank	Between.	Degree	Pagerank
99.934%	99.948%	99.932%	99.97%	99.96%	99.97%	99.93%	99.78%	99.84%

Now we will focus on vertices with high relative centrality. Table 3 lists the id of vertices of the World Series dataset whose relative betweenness centrality is at least 0.5. Some of these nodes compare also in Fig. 4(a). That is, there are nodes that are central in both the static cumulative DRG and in the temporal graphs. Table 4 lists, for all central nodes with respect the betweenness centrality in the World Series temporal graphs (see Fig. 4(a)), the relative centrality in the static cumulative DRG. It is interesting to note that all the listed nodes but one (node 166) belong to the 0.066% (=100 − 99.934, see Table 2) of vertices whose relative centrality is at least 0.01. That is almost all nodes that are central in temporal graphs are also central in the whole graph.

Table 3. Vertices of the whole World Series dataset whose relative betweenness centrality is at least 0.5.

Vertex id	Relative centrality	Vertex id	Relative centrality
299	1.00	122	0.62
31	0.69	11374	0.52
27	0.67		

Table 4. Relative centrality in the whole World Series dataset of nodes that are central in the temporal graphs.

Vertex id	Relative centrality	Vertex id	Relative centrality
299	1.00	243	0.19
31	0.69	340	0.18
27	0.67	126	0.10
122	0.62	516	0.07
14	0.49	521	0.05
46	0.25	66050	0.03
28	0.23	166	< 0.01
523	0.20		

For what concerns the other centrality measures and the Italian Sampling and Black Friday datasets we have similar results that are not reported in this extended abstract for lack of space.

5 Discussion and Conclusions

In this paper we have studied the evolution of four centrality measures (betweenness, degree, closeness, and PageRank) on the DRG temporal retweet graphs based on three datasets: Black Friday, World Series, and Italian Sampling. Our main results can be summarized as follows: (i) too many nodes are central with respect closeness centrality, hence this measure is useless to detect influential users; (ii) for the other measures, the number of nodes with very low centrality is very high and the sets of central nodes (with centrality values above 75% of the maximum) are very small and quite similar in the three measures; (iii) similar results hold also for the static cumulative graphs where the sets of nodes with relevant centrality contain central nodes in the sequence of DRG temporal graphs.

As pointed out in [4], the DRG temporal graphs derived from our datasets are quite sparse: this could explain the small number of central nodes respect to the three centrality measures.

According to the above analysis the approach based on the DRG temporal graph and the centrality measures represent a promising approach for detecting influencer in the microblogging Twitter platform.

References

1. Amati, G., Angelini, S., Bianchi, M., Costantini, L., Marcone, G.: A scalable approach to near real-time sentiment analysis on social networks. In: DART 2014 Information Filtering and Retrieval. Proceedings of the 8th International Workshop on Information Filtering and Retrieval co-located with XIII AI*IA Symposium on Artificial Intelligence (AI*IA 2014), vol. 1314, pp. 12–23. CEUR-WS.org, December 2014
2. Amati, G., Angelini, S., Capri, F., Gambosi, G., Rossi, G., Vocca, P.: Modelling the temporal evolution of the retweet graph. IADIS Int. J. Comput. Sci. Inf. Syst. **11**(2), 19–30 (2016)
3. Amati, G., Angelini, S., Capri, F., Gambosi, G., Rossi, G., Vocca, P.: Twitter temporal evolution analysis: comparing event and topic driven retweet graphs. In: BIGDACI 2016 - Proceedings of the International Conference on Big Data Analytics, Data Mining and Computational Intelligence, Funchal, Madeira, Portugal, July 2–4, 2016, vol. 1 (2016)
4. Amati, G., Angelini, S., Capri, F., Gambosi, G., Rossi, G., Vocca, P.: On the retweet decay of the evolutionary retweet graph. In: Gaggi, O., Manzoni, P., Palazzi, C., Bujari, A., Marquez-Barja, J.M. (eds.) GOODTECHS 2016. LNICSSITE, vol. 195, pp. 243–253. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61949-1_26
5. Bavelas, A.: Communication patterns in task-oriented groups. J. Acoust. Soc. Am. **22**(6), 725–730 (1950)
6. Bonacich, P.: Power and centrality: a family of measures. Am. J. Sociol. **92**(5), 1170–1182 (1987)
7. Borgatti, S.P.: Centrality and network flow. Soc. Netw. **27**(1), 55–71 (2005)
8. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. Comput. Netw. ISDN Syst. **30**(1–7), 107–117 (1998)
9. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, P.K.: Measuring user influence in twitter: the million follower fallacy. ICWSM **10**(10–17), 30 (2010)
10. Conti, M., De Salve, A., Guidi, B., Ricci, L.: Epidemic diffusion of social updates in Dunbar-based DOSN. In: Lopes, L., Žilinskas, J., Costan, A., Cascella, R.G., Kecskemeti, G., Jeannot, E., Cannataro, M., Ricci, L., Benkner, S., Petit, S., Scarano, V., Gracia, J., Hunold, S., Scott, S.L., Lankes, S., Lengauer, C., Carretero, J., Breitbart, J., Alexander, M. (eds.) Euro-Par 2014. LNCS, vol. 8805, pp. 311–322. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-14325-5_27
11. Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry **40**, 35–41 (1977)
12. Freeman, L.C.: Centrality in social networks conceptual clarification. Soc. Netw. **1**(3), 215–239 (1978)
13. Hage, P., Harary, F.: Eccentricity and centrality in networks. Soc. Netw. **17**(1), 57–63 (1995)
14. Kwak, H., Lee, C., Park, H., Moon, S.: What is twitter, a social network or a news media? In: Proceedings of the 19th International Conference on World Wide Web, WWW 2010, New York, NY, USA, pp. 591–600. ACM (2010)

15. Myers, S.A., Sharma, A., Gupta, P., Lin, J.: Information network or social network?: the structure of the twitter follow graph. In: Proceedings of the 23rd International Conference on World Wide Web, WWW 2014 Companion, New York, NY, USA, pp. 493–498. ACM (2014)
16. Nieminen, J.: On centrality in a graph. Scand. J. Psychol. **15**, 322–336 (1974)
17. Shimbel, A.: Structural parameters of communication networks. Bull. Math. Biophys. **15**(4), 501–507 (1953)

Dynamics of Emotions and Relations in a Facebook Group of Patients with Hidradenitis Suppurativa

Gianfranco Lombardo, Alberto Ferrari, Paolo Fornacciari, Monica Mordonini, Laura Sani, and Michele Tomaiuolo(✉)

Dipartimento di Ingegneria e Architettura, Università di Parma, Parma, Italy
michele.tomaiuolo@unipr.it

Abstract. Hidradenitis suppurativa (HS) is an orphan, underdiagnosed and painful disease of the skin that has a considerable negative impact on quality of life and on emotional well-being. As reported by the italian HS patients' association (Inversa Onlus), this condition brings patients to develop an emotional closure with the consequence that they often don't talk about their condition with anybody. In this paper we discuss some results obtained by applying automatic emotion detection and social network analysis techniques on the Facebook group of the Inversa Onlus association. In particular, we analyze the patients' emotional states, as expressed by the post published from 2010 to 2016, and how these emotions are influenced by friendships in the group, during the years.

Keywords: Social network analysis · Emotion detection
Sentiment analysis · Hidradenitis suppurativa · Facebook

1 Introduction

Hidradenitis suppurativa (HS) is an orphan and underdiagnosed disease of the skin that affects 1%–4% of the general population. The main characteristics of HS are chronic, inflammatory, painful boils in the folds of the skin that have a considerable negative impact on quality of life and on emotional well-being. It is widely reported that the difficulties in obtaining a diagnosis together with the lack of proper therapies bring HS patients to develop an emotional closure, with the consequence that patients often don't talk about their condition with anybody, as reported by the italian HS patients'association (Inversa Onlus) [1]. In the last few years, with the diffusion of social media, a considerable number of patients started to share a considerable amount of data related to their feelings. In this paper we present the results obtained using Sentiment Analysis on Inversa Onlus's Facebook group for the automatic detection of the emotional state of the patients from 2010 to 2016. In particular, this work focuses on the analysis of how friendships in the group influence the emotions expressed by patients

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 269–278, 2018.
https://doi.org/10.1007/978-3-319-76111-4_27

over the years. Results have been obtained using a manual annotated training set and a seven-output hierarchical classifier system based on Parrot's emotion categorization (joy, love, surprise, fear, anger, sadness, and objective content) [2].

2 Related Works

Nowadays, online social networks provide more and more medical data; patients often share personal clinical information within an online community, with the aim of receiving emotional support. In recent years, the interest in patients' opinions and feelings expressed in web communities has considerably increased. One of the biggest challenges is to get a clear understanding of patients' condition. In [3] authors identify the 15 largest Facebook groups focused on diabetes management, analyzing 690 comments from wall posts written by 480 unique users. The work aims at identifying, with traditional manual method of content analysis, the main topics of the discussions. Since this approach is not scalable when trying to analyze huge quantities of data, it is necessary to introduce automatic analysis, based on machine learning algorithms. In [4] authors analyze different sources of patients' information (social media, blog, patients networks) in order to detect poor quality healthcare using sentiment analysis and natural language processing. Sentiment analysis using machine learning algorithms represents an automatic way to analyze sentiments, emotions and opinions from written language [5–7] and it's becoming increasingly important in the context of social media and for business and social sectors [8]. Our approach is innovative in joining two types of analysis:

- emotion detection applied to patients' Facebook posts, based on a hierarchical classifier using Parrot's emotion categorization (joy, love, surprise, fear, anger, sadness, and objective content) [2];
- analysis of the Facebook friendship relations among the members using Social Network Analysis, a discipline that focuses on the structural and topological features of the network, that can be useful in order to identify models of participation, opinion leaders and behaviors adoption [9].

As for the tools used, hierarchical classification is widely applied to big data collections [10–12] and in the scientific literature it is often considered an effective approach for emotion detection [13, 14].

3 Methodology

In this section we present all the tools and algorithms we used in order to analyze and assign an emotion to the posts of each patient. According to Parrott's socio-psychological model, all human feelings could be partitioned with a primary categorization composed of six emotions, three positive (joy, love, surprise) and three negative (fear, sadness, anger). In order to obtain a performing system for emotion detection we used stopwords and stemming algorithms to preprocess the data, the bag-of-words model to extract features from the sentences, the Information Gain algorithm in order to reduce the features, and the Naive Bayes Multinomial classifier to optimize some hyperparameters in order to achieve better accuracy [15]. In fact, the whole system is built after a selection of the best mechanisms and parameters, in accordance to previous experiences [13]. Each step is described below, and the execution flow is presented in Fig. 1.

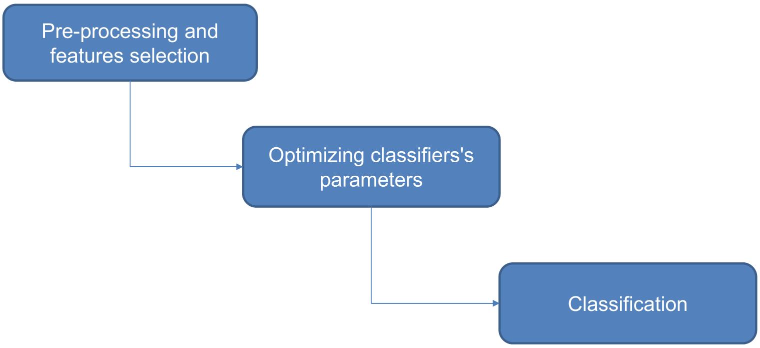


Fig. 1. Execution flow

3.1 Creating the Training Set

The first problem that we have considered has been the creation of a performing training set, to classify the patients' emotions. Due to the lack of useful datasets containing annotated posts of Italian patients, we have decided to create our own training set based on supervised learning and manual annotations. For creating the training set, we have considered the 10% of all the available posts, published in the Facebook group in the past seven years (Table 1).

Table 1. Number of posts in the training set for each emotion.

Sentiment	Number of posts in the training set
Objective	85
Love	70
Joy	80
Surprise	30
Fear	75
Sadness	85
Anger	65

3.2 Pre-processing and Features Selection

All posts have been pre-processed, in order to preserve only the elements with an emotional meaning. For this reason, we have used different automatic filters to remove Italian stopwords and punctuation, encode special characters, correct spelling mistakes, substitute contractions with their textual extension, and substitute smiles and emoticons with appropriate words. At the end of this process, sentences have been also filtered with a Stemming algorithm, in order to reduce inflected words to their word stem.

The last step has been the generation of features based on bag-of-words model using the StringToWordVector algorithm, turning each string into a set of attributes representing word occurrences. The TF-IDF function has been used to evaluate the relevance of each word in each sentence. Finally, these attributes have been filtered using the Information Gain algorithm, in order to extract the most meaningful features.

3.3 Classification

In this subsection we present the approach that we have used to build the resulting classifier. With the aim of building a hierarchical classifier, the manually annotated training set has been used to create four different training sets, in which each sentence is labeled according to the task of each classifier (Fig. 2):

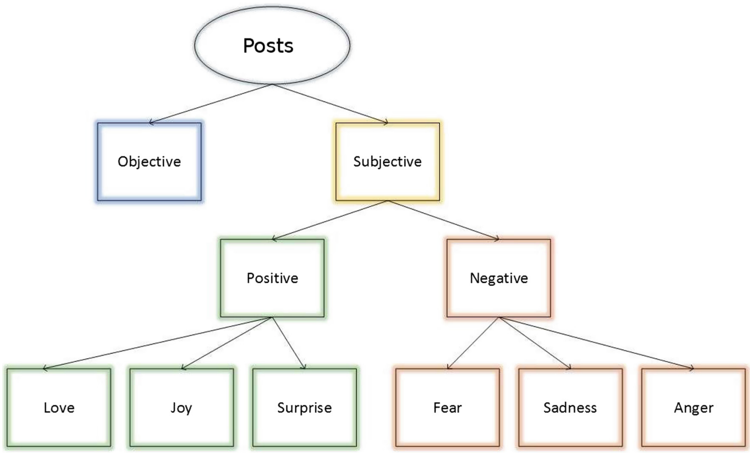


Fig. 2. Hierarchical classifier used

– **Objectivity/subjectivity classifier:** The training set used is simply the pre-processed training set in which every post associated with an emotion has been labelled as subjective while the remaining others, for example information requests or communications from the patient's association, have been labelled as objective.

- **Polarity classifier:** The training set used is the pre processed training set in which all of the posts previous labelled as subjective are divided into positive and negative.
- **Positive classifier:** The training set used is composed of the previous posts labelled as positive, divided into love, joy and surprise.
- **Negative classifier:** The training set used is composed of the previous posts labelled as negative, divided into anger, fear and sadness.

3.4 Optimizing Classifiers's Parameters

For each classifier used to build the seven-output hierarchical classification system, a preliminary analysis was performed. In this phase, we have selected the Naive Bayes Multinomial algorithm for implementing each classifier, as it produces the best results. Moreover, we have optimized some parameters, considered relevant for the training phase. In particular, we have searched at the same time for (*i*) the optimal length of N-grams to be used as features, and (*ii*) the number of features to select through the Information Gain algorithm. We have used a grid search, which is simply an exhaustive search through a manually specified subset of the parameters hyperspace of a learning algorithm, in order to select a grid of configurations using cross-validation to estimate the quality of classifiers configured according to them. Results are shown in Table 2.

Table 2. Parameters optimization results

Classifier	N-Gram (max)	Features
Sub/Obj	3	100
Pos/Neg	2	250
Fear/Sadness/Anger	3	340
Love/Joy/Surprise	2	100

4 Results

After having classified the emotions in patients' posts, in a period of seven years, we have performed further analysis. In this section we present the results, which may provide valuable hints for better understanding the Hidradenitis Suppurativa disease and its impact on patients' lives.

4.1 Distribution of Emotions

In this subsection we present the results obtained by analyzing the distribution of emotions for each month of the analyzed year. Looking at the distributions for each year (Fig. 3), and in particular the average distribution (Fig. 4), it is possible to note that the month in which patients are most active in the group is usually September, and to a lesser but considerable extent, January and April-May. These results could be read in different ways, but these three time frames

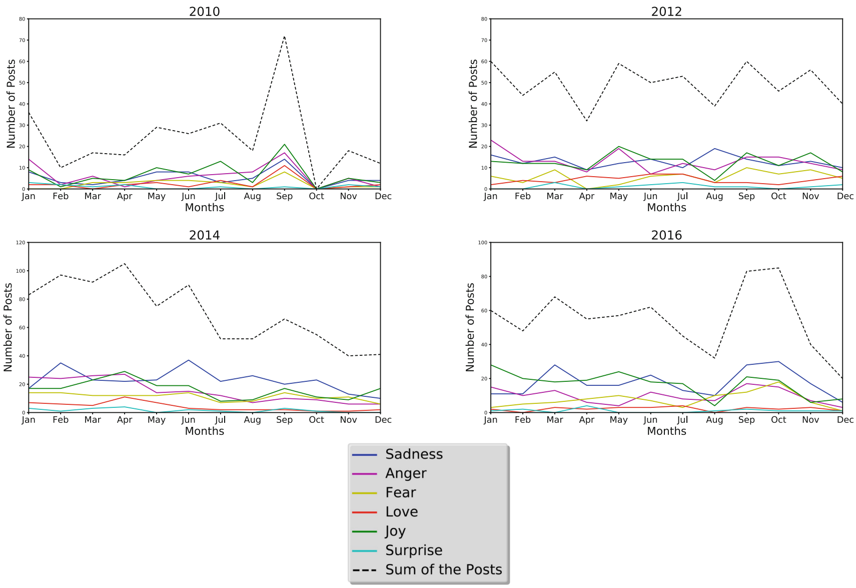


Fig. 3. Distribution of emotions in 2010, 2012, 2014, 2016.

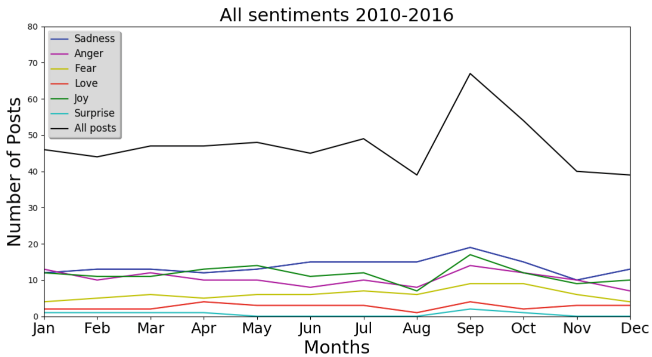


Fig. 4. Average distribution of emotions between 2010–2016.

have in common the fact that they follow holidays. In particular, the most active month (September) follows the traditionally long summer holidays. According to the negative impact of this disease [16], the prevailing emotion is sadness. Summing up the posts conveying negative emotions, they constitute the major component of the available posts. Despite that, the second prevailing emotion is joy, and in particular at the beginning of the group in 2010 it was the most present. This juxtaposition could be read in different ways, but it seems that in the group there are some influencers which express positive emotions in reaction to other negative posts. We will discuss this aspect later, analyzing the results concerning the social network analysis.

4.2 Social Network Analysis

After retrieving information about friendships on Facebook among the group members and classifying posts written by each patient in the group for each year, it is possible to analyze the established social network in the group and its evolution over the years (Fig. 5). In particular, at the beginning of the group activities, in 2010, there were few members, with many Facebook friendship relationships, expressing a variety of emotions. During the following years, the number of members in the group has constantly increased. At the same time, the differences about the emotions expressed by isolated members and connected members have grown.

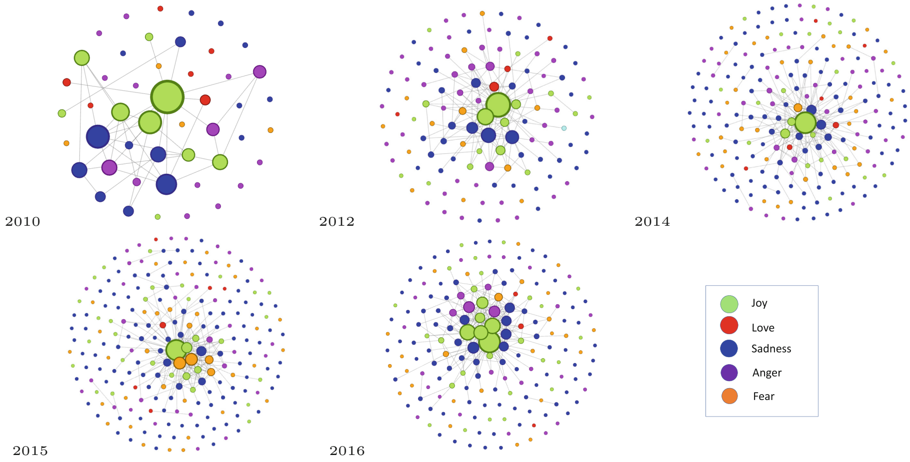


Fig. 5. Social network in 2010, 2012, 2014, 2015, 2016.

4.3 Friendship Relations and Expressed Emotions

The first result presented in this section is obtained analyzing how the emotions expressed by patients are correlated with their own degree in the social network (i.e., the number of friends in the group). Members are classified in four categories, according to the number of friendship relations in the group:

- **Zero Relationships**
- **Weak:** from 1 to 5 Relationships
- **Moderate:** from 6 to 19 Relationships
- **Strong:** 20 or more relationships.

The results obtained analyzing posts and social network in 2016 and in the period between 2010 and 2016 are presented in Fig. 6. It is worth noting that, for the nodes with higher degree, the predominant emotion is joy, while all the negative emotions decrease. As a possible interpretation, this result may represent an evidence of the positive influence of the group. Figure 6, for example, shows that all the patients that have developed many relationships with other peers in the group express positive emotions.

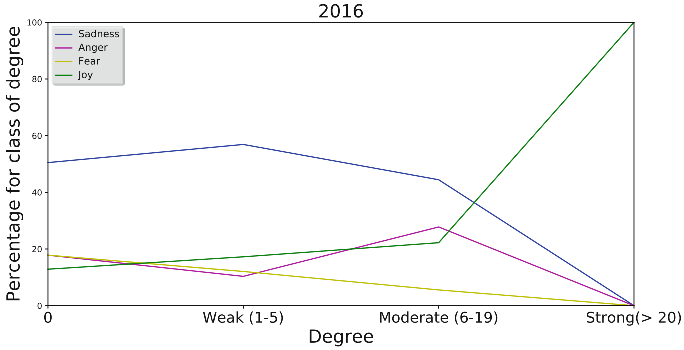


Fig. 6. Relationships between number of friendships and feelings expressed only in 2016

4.4 Changes of Emotions During the Years

Looking at the network between 2014 and 2016, it is clear that the major part of the active members don't have relationships with each others and they express mostly negative emotions. Despite that, analyzing in an anonymous way each patient, we have discovered that, in particular in the period from 2014 to 2016, members that used to express fear or anger emotions, but that were in connection with others, started to develop and express joy emotions. This result has been further analyzed, by calculating a matrix of transitions among emotions, considering the changes of emotions expressed by members with at least three relationships inside the group (Table 3). Observing the matrix, it is possible to note that transitions from anger or fear emotions to joy are more frequent than the transitions from sadness to joy. This result is important to understand that angry and worried members are more easily influenced by the rest of the social network. Instead, sadness appears to be the most static emotion. This essentially means that members who express sadness continue to express sadness, regardless of their own friendships in the group. At the same time, positive emotions as a whole are the most static class of emotions. In fact, transitions from positive to positive emotions are more frequent (53%) than transitions from positive to negative emotions (47%).

According to this multi-faceted analysis, it is thus possible to observe a significantly positive evolution of the group. However, while these analytical results were quite hoped by the group representatives, they were not obvious before this work. Thus, they also highlight the importance of the research and development of this kind of software tools. In future, it would be important to verify if this kind of social and psychological dynamics is common to similar online communities. In fact, to compare more data and possibly generalize our understanding of the phenomenon, we plan to continue this line of research and perform similar analysis in other online groups of patients, of different diseases.

Table 3. Matrix of transitions among emotions

	Joy	Love	Anger	Fear	Sadness
Joy	48%	11%	15%	4%	22%
Love	29%	0%	29%	13%	29%
Anger	29%	7%	14%	14%	36%
Fear	25%	17%	17%	8%	33%
Sadness	18%	3%	12%	6%	61%

5 Conclusion

The advent of social networks has led to new ways of expression for a support group composed of people who share the same disease or trouble. To date, it is still an open challenge of technology to provide tools suitable for a deepened analysis of these online communities and their influence over their participants. In this paper we have reported the results of an investigation of the dynamics of emotions and relations in a Facebook group of patients with hidradenitis suppurativa, who have been studied from 2010 to 2016.

We have detected the emotions of the members' posts by applying a hierarchical classification system, and then we have calculated the emotion of each member of the group and correlated it with his/her degree of relations in the group, that is the number of his/her friendship relations with other members of the community. This type of analysis has been performed for each year of the considered period.

The results have shown some interesting correlations among the decrease of negative emotions and the increase of the degree of relations among members. More generally, as the time progresses, it is possible to observe a positive dynamics of the members' emotions.

These results, though hoped, were not so obvious to the group representatives. In our opinion, such results highlight the importance of the research and development of this kind of software tools. In future researches, we plan to perform a similar analysis in other online groups of patients, of different diseases. This way, we will be able to compare the different results and verify if this kind of social and psychological dynamics is common to such online communities.

References

1. Bettoli, V., Pasquinucci, S., Caracciolo, S., Piccolo, D., Cazzaniga, S., Fantini, F., Binello, L., Pintori, G., Naldi, L.: The hidradenitis suppurativa patient journey in Italy: current status, unmet needs and opportunities. J. Eur. Acad. Dermatol. Venereol. **30**, 1965–1970 (2016)

2. Parrott, W.G.: Emotions In Social Psychology: Essential Readings. Psychology Press, Hove (2001)
3. Greene, J.A., Choudhry, N.K., Kilabuk, E., Shrank, W.H.: Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J. Gen. Intern. Med. **26**, 287–292 (2011)
4. Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., Donaldson, L.: Harnessing the cloud of patient experience: using social media to detect poor quality healthcare. BMJ Qual. Saf. **22**, 251–255 (2013)
5. Liu, B.: Sentiment analysis and opinion mining. Synth. Lect. Hum. Lang. Technol. **5**, 1–167 (2012)
6. Mohammad, S.M.: Sentiment analysis: detecting valence, emotions, and other affectual states from text. In: Emotion Measurement, pp. 201–238 (2015)
7. Kao, E.C.C., Liu, C.C., Yang, T.H., Hsieh, C.T., Soo, V.W.: Towards text-based emotion detection a survey and possible improvements. In: International Conference on Information Management and Engineering, 2009. ICIME 2009, pp. 70–74. IEEE (2009)
8. Franchi, E., Poggi, A., Tomaiuolo, M.: Social media for online collaboration in firms and organizations. Int. J. Inf. Syst. Model. Des. (IJISMD) **7**, 18–31 (2016)
9. Angiani, G., Fornacciari, P., Iotti, E., Mordonini, M., Tomaiuolo, M.: Models of participation in social networks. In: Social Media Performance Evaluation and Success Measurements, p. 196 (2016)
10. Dumais, S., Chen, H.: Hierarchical classification of web content. In: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 256–263. ACM (2000)
11. Silla Jr., C.N., Freitas, A.A.: A survey of hierarchical classification across different application domains. Data Min. Knowl. Discov. **22**, 31–72 (2011)
12. Addis, A., Armano, G., Vargiu, E.: A progressive filtering approach to hierarchical text categorization. Commun. SIWN **5**, 28–32 (2008)
13. Angiani, G., Cagnoni, S., Chuzhikova, N., Fornacciari, P., Mordonini, M., Tomaiuolo, M.: Flat and hierarchical classifiers for detecting emotion in tweets. In: Adorni, G., Cagnoni, S., Gori, M., Maratea, M. (eds.) AI*IA 2016. LNCS (LNAI), vol. 10037, pp. 51–64. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49130-1_5
14. Ghazi, D., Inkpen, D., Szpakowicz, S.: Hierarchical approach to emotion recognition and classification in texts. In: Farzindar, A., Kešelj, V. (eds.) AI 2010. LNCS (LNAI), vol. 6085, pp. 40–50. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13059-5_7
15. Chen, J., Huang, H., Tian, S., Qu, Y.: Feature selection for text classification with naïve bayes. Expert Syst. Appl. **36**, 5432–5435 (2009)
16. Onderdijk, A., Van der Zee, H., Esmann, S., Lophaven, S., Dufour, D., Jemec, G., Boer, J.: Depression in patients with hidradenitis suppurativa. J. Eur. Acad. Dermatol. Venereol. **27**, 473–478 (2013)

OOPP: Tame the Design of Simple Object-Oriented Applications with Graphical Blocks

Alberto Ferrari, Gianfranco Lombardo, Monica Mordonini, Agostino Poggi, and Michele Tomaiuolo(✉)

Dipartimento di Ingegneria e Architettura, Università di Parma, Parma, Italy
{alberto.ferrari,gianfranco.lombardo,monica.mordonini, agostino.poggi,michele.tomaiuolo}@unipr.it

Abstract. Many and varied experiences are being reported, about the first introduction to programming for young students and neophytes. However, tools and methodologies are needed also for a more comprehensive learning process, which requires to design the architecture of any small but functioning application. We propose a new environment, based on the use of graphical blocks, for designing some object-oriented applications. It merges the positive features of block-programming with the object-oriented paradigm in a graphical educational environment. It is developed as a tool for supporting the objects-early approach. The whole methodology is targeted at high school students, university freshmen and unemployed people who are motivated to learn to code professionally. In these cases, where we have firstly experimented this approach, the concepts of object-oriented programming (OOP) cannot be relegated to a secondary role, but they have to be introduced early and presented in their most intuitive form.

Keywords: Computer programming · Block programming
Object-oriented programming · Education

1 Introduction

Most recent reports, regarding the introduction to coding, target a young or a very young audience [4,17,20]. A feature shared by a number of these projects is the use of puzzle programming for simplifying the very first approach to computer programming, trying to eliminate the syntactic burden.

Our work is oriented in particular to programming courses for high school students and university freshmen. Moreover, there is a growing social pressure to reorient unemployed people towards computer programming, through vocational training courses. In all these cases, the typical syntactic difficulties related with coding are soon accompanied by some challenges of application design. In fact, courses for professional computer programming cannot procrastinate the concepts of object orientation, which become harder to learn at a later stage. Thus, we are working to define an approach which can get the best of both the

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 279–288, 2018.
https://doi.org/10.1007/978-3-319-76111-4_28

worlds of object-oriented programming and puzzle programming. In this approach, which we call *Object-Oriented Puzzle Programming (OOPP)*, we propose a tool for the design and development of object-oriented applications, adopting the objects-early methodology [15]. Applications are realized by connecting visual objects representing the fundamental elements of object-oriented programming, in a similar manner to puzzle programming. The user/programmer realizes a puzzle of connected blocks, which represents the object-oriented structure of the application.

The rest of the article is organized in the following way. Section 2 presents the current discussion and the situation about the introduction to coding, especially for high school students and young workers. Section 3 discusses the methodology and tool we propose for Object Oriented Puzzle Programming. Section 4 presents an evaluation of the tool by a class of post-high school students. Finally, some concluding remarks are presented.

2 Related Works

The whole idea of Computational Thinking was popularized by Wing [21], as "*a way of solving problems, designing systems, and understanding human behavior that draws on concepts fundamental to computer science.*" Apart from coding, other aspects and concepts are fundamental to Computational Thinking [6,7], including abstraction and modeling.

The founding idea of block programming, which characterize a number of projects [4,16,22] related to the introduction to Computational Thinking, is to provide a graphical interface with blocks of diverse types. The user/programmer can combine the blocks in various ways through drag and drop operations, in such a way to form a puzzle representing the solution to a given problem. The blocks represent the basic elements of the language and they are differentiated by form and color, to ease their identification and usage. The difficulties, faced by novices which approach programming for the first time [1,3,9], are essentially of two different kinds: on the one hand it is necessary to formalize the ideas about the solution in the form of an algorithm, on the other hand it is necessary to adhere to the rigorous syntax of a programming language. Block programming eliminates the possibility to introduce syntactic errors and allows users to focus entirely on the logics of assigned problems and their solutions. In fact, the composition of blocks is rigidly constrained by existing slots, which represent the syntactic constraints of the language.

Although most educational projects and experiences proposing a block programming approach are oriented to a very young audience, projects directed toward high school students are beginning to appear. In [19] some points of strength and weakness are analyzed, as they are perceived by high school students which are allowed to move from a textual programming environment to a block-based one.

In the last decades, the object-oriented approach has become one of the most adopted programming paradigm [12,13]. It is popular both at the educational and professional level. This fact is largely accepted, but there are still a lot of debates focusing on the time in which OOP is best introduced in CS1 courses. The debate about the "paradigm shift" is still actual, but in the last years in many introductory university courses there has been a move from the procedural paradigm to the object-oriented one. A comparative analysis [18] shows that student who have started their curriculum with an object-first approach obtain better results, when they have to design software for solving complex problems. Other studies [2,5,10] present experiences in which the object-first approach has effectively led to great improvements. They show that students obtain better overall results and the failure rate is drastically reduced. An additional advantage of the object-first approach is to facilitate the comprehension of the object-oriented paradigm, while the shift from a paradigm to another is quite difficult for those who started studying the discipline with an imperative/procedural approach. In [12] it is argued that "*it is the switch that is difficult, not object-orientation.*" If the main difficulty lies in the paradigm-shift, from the procedural paradigm to the object-oriented one, then the objects-early approach has the lowest difficulty level, as can be seen in many CS1 courses.

In introductory courses based on the object-early approach, professional tools can be too complex, especially for newbies. Among tools tailored for educational purposes, instead, BlueJ[1] provides a useful IDE for object-early didactics. However, unlike the tool we are going to present, there is little functionality for designing the internal structure of the classes of an object-oriented application.

3 Object Oriented Puzzle Programming

OOPP is an integrated development environment primarily designed for teaching purposes in order to introduce object-oriented design and programming in a simple and intuitive way. OOPP is a tool that allows a user/programmer to design and create simple object-oriented applications by block composition, and then automatically generate the target code that will be the frame of the application to be realized. It includes a workspace that allows you to define, in a visual way, "blocks" of code, with related methods, constructors and attributes, that will then automatically generate the corresponding target code.

In Blockly[2], we have developed a set of high-level blocks that can be used for object-oriented programming. The visual environment is developed entirely in JavaScript and therefore it can be executed within any web browser. This fact further simplifies the use of the application, that does not require any specific software installation. The environment is characterized by three main elements (Fig. 1):

1. a toolbox which contains the available blocks, organized by their type;
2. a work space where to place and link the blocks to form a puzzle/program;
3. a text box to display the corresponding generated code automatically, by converting each block of the puzzle into a sequence of target instructions.

[1] https://www.bluej.org/.
[2] https://developers.google.com/blockly/.

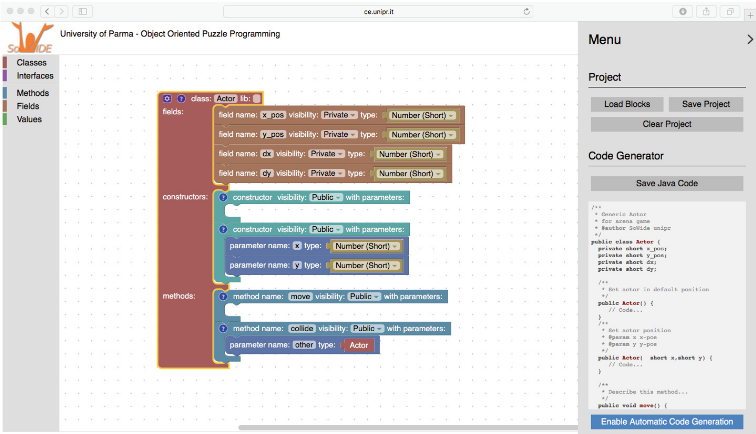


Fig. 1. Object-Oriented Puzzle Programming environment.

3.1 The Blockly Framework

Blockly includes some standard categories of blocks: logical blocks, cycles, math operators, lists, variables, and so on. Starting from these categories, we have created additional blocks for object-oriented programming: interfaces, classes, constructors, methods, and parameters. The new blocks are shaped to allow only connections that respect the syntactic constraints. Defining a new class or interface makes a new block in the toolbox instantly available, which is the new type of data that you can then use in developing the application. To keep the environment simple and intuitive, as much as possible, only the main features of object-oriented programming have been introduced. For example, it is not possible to explicitly define inner classes and abstract classes. With the present features, it is possible to put into practice the fundamental concepts of object-oriented programming such as encapsulation, inheritance, and polymorphism. Our didactic experience shows us that more advanced features, which can be found for example in the UML language specifications, are neither necessary nor desirable in such an introductory environment. OOPP is thought to be a tool for the initial stages of a didactic path for object-oriented design. Hence, it is unsuitable for complex applications. As the same authors of Blockly say: "*Blockly is currently designed for creating relatively small scripts... Please do not attempt to maintain the Linux kernel using Blockly.*" – Neil Fraser - Google.

Blockly is defined as "*a library for building visual programming editors*", that is a tool for application developers, and many are in fact the educational applications that use it as a starting point for development. It provides a workspace that allows users (novice programmers) to write programs by linking the various blocks. Each visual object (block) actually represents a code object. Blockly also allows to create new blocks. Then, these new blocks can be used in a variety of applications. In a sense, creating a set of new blocks corresponds to the creation of a new language with its new syntax and semantics. The syntax is defined by the structure of the blocks, their color, their shape and the possible connections. In practice, Blockly allows to write only syntactically correct programs, with textual code replaced by visual blocks. The *Block Factory* (Fig. 2) is a graphical tool of Blockly, which can be used to start creating new blocks, by defining their main features through a simple and intuitive interface. Block semantics is then defined by providing the corresponding code for each of them, which can be written in any programming language. The double representation of the program (Fig. 1), both as a puzzle and as a textual code, is very interesting and stimulating in an educational setting. In fact, the novice programmer can compare the two versions.

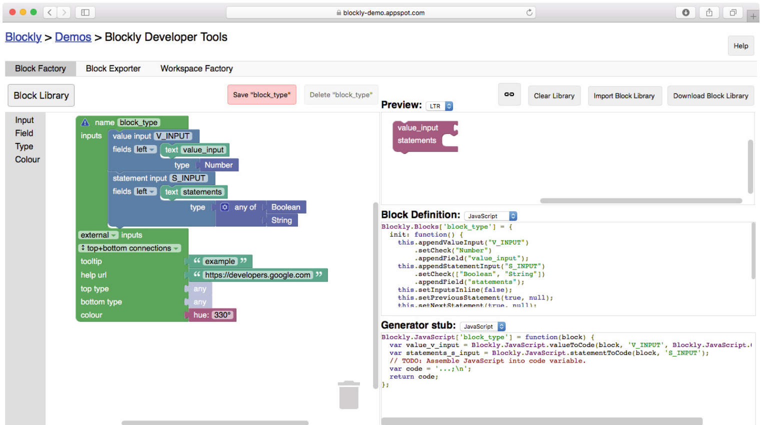


Fig. 2. Blockly Block factory.

3.2 OOPP: The Working Environment

OOPP is a web application developed in JavaScript and entirely executed within a browser. Its graphical interface provides a toolbox that presents the various categories of blocks that can be dragged and linked together to form a puzzle in a work area. Moreover, it provides a context menu that allows the activation of all the features inherent to the current project and necessary for the generation of code. A puzzle representing the set of interfaces and classes of an application is defined by a set of blocks of various types. In particular, we have extended Blockly with object-oriented blocks. Besides the blocks that define interfaces and classes, there are other blocks for attributes, constructors, methods, parameters, data types, etc. An entire OOPP project can be saved into a file, and then loaded again.

OOPP has been designed with the main goal of easing its use to people with limited or no knowledge on object-oriented programming. Therefore, the information to manage during the definition of a puzzle has been minimized. In each particular phase of the definition, it shows only the components that

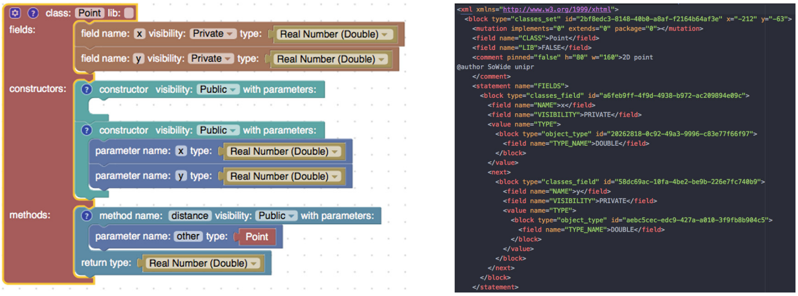


Fig. 3. A simple class and part of its internal representation.

are considered necessary. For example, inheritance is not proposed upfront when starting to define a new class (Fig. 3).

When an application involves a large number of blocks, usually it is necessary to switch from a global view, which allows the analysis of the structure of the application, to a specific view, which allows the analysis of the relationships among a subset of the blocks of the application, or the analysis, modification and extension of a single block. To do it, OOPP allows the visualization at different zoom levels, providing a "collapse" and an "expand" command to automatically minimize and recover to the original size one or all the blocks of the application. For classes and interfaces, it is possible to mark them as provided by an external library; this way, only an *import* instruction is generated.

3.3 Java-Based Projects

The first target language supported by OOPP is Java [8]. Each time that the puzzle is modified in any way, the corresponding Java code is automatically generated anew and shown. This real-time update of the code has the aim to highlight the syntactic features of the target language. Students generally appreciate it a lot, since it allows them to compare the generated code with the block structure, which appears clearer. However, the automatic generation of code can be temporarily disabled. This option is suggested to avoid slowing down the application, in case of particularly complex projects, or while loading large libraries.

The translation of blocks into code is specified in JavaScript files, created for the peculiar blocks of this project, about interfaces, classes, and operational Java code. The constraints, which limit the possible connections among blocks, eliminate most syntactic errors. However, more checks have been implemented, to avoid other common syntactic and semantic errors. This aspect is very important, since the application has an educational scope; as it is documented also in other studies [11], syntactic errors and other issues, related to imprecisely typed code, are one of the most important cause of demoralization and discouragement for neophytes of coding. As an example, we cite here a quite classical error, which is not merely syntactic, in this kind of projects: a method is not implemented in a class, even if some of its interfaces specify it.

As we have seen, using *mutators* it is possible to organize classes and interfaces in a complex hierarchy. For example, in an application it is possible to define a class and associate it with a certain interface. Figure 4 shows a puzzle representing this example: class X implements interface Y, without providing an implementation of its method Z. From the point of view of blocks, no constraints are violated. However, while generating the corresponding Java code, the process cannot complete and an explicative message for the error is shown. The analysis is performed by collecting all signatures of methods defined in implemented interfaces, and confronting them with methods defined in the class itself; only in the case they correspond, the Java code is generated.

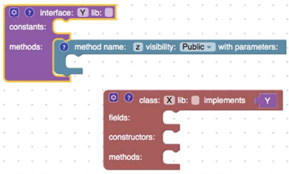


Fig. 4. A simple class and relative interface.

In association with the development environment, an extension has been developed, which allows users to import class and interfaces with all their features. Analyzing a Jar file, the tool can generate an OOPP project file, containing blocks corresponding to all defined classes and interfaces. Various options are available, for example to let blocks appear in "collapsed" mode when loaded, or to set classes and interfaces as "lib", and thus not associated with operational code, but only generating *import* statements. The application uses Java reflection and it can generate a new OOPP project, even when Java source files are not available. Since it is mainly intended as a didactic tool, the import extension does not handle the most complicated settings, e.g. references to external classes, etc.

4 Evaluation

The use of OOPP has been experimented in an introductory CS course, in a school for helping young unemployed students to qualify for developer jobs (an Italian "Corso di Alta Formazione"). The course is organized along the principles of the objects-early methodology, following few introductory lessons about imperative programming. The main objective of the course is to introduce the basic concepts of object-oriented programming and object-oriented software application design, through the Java language. More than 60% of students (Fig. 5), before the course, had no skills or competences about computer programming, at all. An additional 16,7% had not developed any object-oriented application before, though knowing the paradigm from the theoretical point of view.

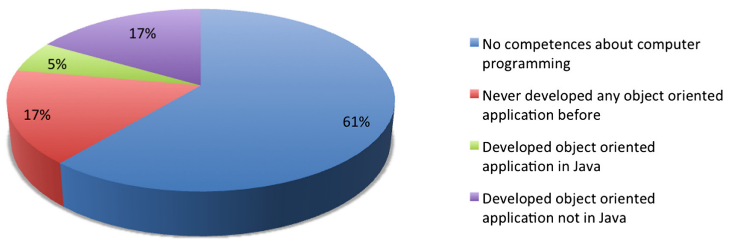


Fig. 5. Competences about computer programming.

In the very first part of the course, lab activities have been characterised by the use of objects of predefined classes, and their reciprocal interaction. This educational approach is inspired by the ideas presented by Kölling [14] and by the BlueJ development environment.

At the moment of designing and realizing classes, and when introducing the concepts of inheritance and polymorphism, students have had the choice to use OOPP, or a more traditional environment. In the first phase, all students have experimented the use of OOPP to design their application and generate the structure of classes for their applications.

Almost all students have appreciated the usability of the OOPP environment; also among neophytes of object-oriented programming, more than 90% has evaluated OOPP as simple to use (Fig. 6). It has been found useful, at least for designing classes for various proposed problems, by 67% of students; 28% of them designed classes for all problems using OOPP environment. In particular, students have reported that, operating through the composition of blocks, it is easier both (i) to define the various needed elements (fields, methods, constructors) of a class, and (ii) to understand the syntactic structures of the generated code, which is available in real-time during the composition of the puzzle.

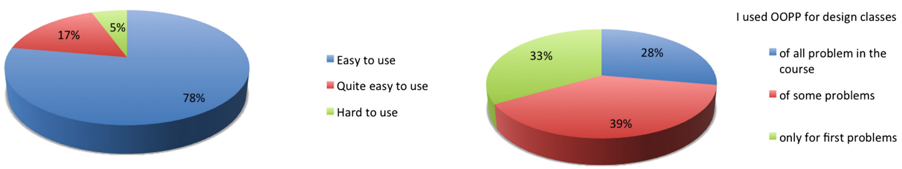


Fig. 6. Usability and use of OOPP.

4.1 Future Developments

The system functionalities have been experimented successfully, for the part related to design, i.e., it allows users to define the high level class model of a Java application in a simple and intuitive way.

A possible development regards the possibility to generate code in more programming languages, implementing the object-oriented paradigm. In particular, one underway extension regards the Python language, in the light of its growing usage in educational contexts. Apart from minor adjustments to blocks, different management of interfaces and support for multiple inheritance, it is necessary to redefine block-code conversions. Anyway, the application structure remains almost the same.

Though not being the initial scope of this project, an extension can also be thought, for allowing a user to develop a complete application, all inside a unified OOPP project. For constructors and methods, it is necessary to provide an implementation, using the operational blocks already available with other systems based on Blockly, together with specific blocks for instantiation and usage of objects. In this case, the comprehensive puzzle obviously becomes larger, or much larger. Thus, the use of such operational blocks inside a unified OOPP project is suggested only for very simple applications. However, the `zoom` and `collapse` functionalities can help to manage a workbench with plenty of blocks.

5 Conclusions

In this paper, we have presented OOPP, a project that merges puzzle programming and object-oriented paradigm in the form of a didactic tool to support the objects-early methodology for introducing the object-oriented paradigm, for high school students and university freshmen. OOPP is a web application which provides blocks for the basic elements of object-oriented programming, i.e., classes, attributes and methods. A user/programmer can create a puzzle as a collection of connected graphical programming blocks. The puzzle could then be automatically translated to Java or any other supported language. In fact, OOPP is not tied up to a specific target language. We think that our project will be useful in CS1 courses, but also in vocational courses, as a starting environment for modeling and designing classes. Automatic generation of code from a puzzle of blocks to diverse object-oriented languages is a valid help to compare syntactical differences from a unique starting project.

References

1. AICA: Informatica nei licei nel contesto della riforma della scuola, documenti di Mondo Digitale (2003)
2. Bennedsen, J., Schulte, C.: What does objects-first mean?: an international study of teachers' perceptions of objects-first. In: Proceedings of the Seventh Baltic Sea Conference on Computing Education Research, vol. 88, pp. 21–29. Australian Computer Society, Inc. (2007)

3. CINI, GII, GRIN: Manifesto sull'informatica nella riforma della scuola superiore. http://www.grin-informatica.it/opencms/export/sites/default/grin/files/manifesto.pdf
4. Cooper, S., Dann, W., Pausch, R.: Alice: a 3-D tool for introductory programming concepts. J. Comput. Sci. Coll. **15**(5), 107–116 (2000)
5. Cooper, S., Dann, W., Pausch, R.: Teaching objects-first in introductory computer science. ACM SIGCSE Bull. **35**(1), 191–195 (2003)
6. Denning, P.J.: The profession of it beyond computational thinking. Commun. ACM **52**(6), 28–30 (2009)
7. Denning, P.J., Comer, D.E., Gries, D., Mulder, M.C., Tucker, A., Turner, A.J., Young, P.R.: Computing as a discipline. Computer **22**(2), 63–70 (1989)
8. Ferrari, A., Poggi, A., Tomaiuolo, M.: Object oriented puzzle programming. In: Didattica Informatica-Didamatica 2016, pp. 1–10 (2016)
9. Gander, W., Petit, A., Berry, G., Demo, B., Vahrenhold, J., McGettrick, A., Boyle, R., Mendelson, A., Stephenson, C., Ghezzi, C., et al.: Informatics education: Europe cannot afford to miss the boat. ACM (2013). http://europe.acm.org/iereport/ie.html
10. Gries, D.: A principled approach to teaching OO first. SIGCSE Bull. **40**(1), 31–35 (2008). http://doi.acm.org/10.1145/1352322.1352149
11. Kelleher, C., Pausch, R.: Lowering the barriers to programming: a taxonomy of programming environments and languages for novice programmers. ACM Comput. Surv. (CSUR) **37**(2), 83–137 (2005)
12. Kölling, M.: The problem of teaching object-oriented programming, part 1: languages. J. Object-Oriented Program. **11**(8), 8–15 (1999)
13. Kölling, M.: The problem of teaching object-oriented programming, part 2: environments. J. Object-Oriented Program. **11**(9), 6–12 (1999)
14. Kölling, M.: Using BlueJ to introduce programming. In: Bennedsen, J., Caspersen, M.E., Kölling, M. (eds.) Reflections on the Teaching of Programming: Methods and Implementations. LNCS, vol. 4821, pp. 98–115. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-77934-6_9
15. Lister, R., Berglund, A., Clear, T., Bergin, J., Garvin-Doxas, K., Hanks, B., Hitchner, L., Luxton-Reilly, A., Sanders, K., Schulte, C., Whalley, J.L.: Research perspectives on the objects-early debate. SIGCSE Bull. **38**(4), 146–165 (2006). http://doi.acm.org/10.1145/1189136.1189183
16. Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., et al.: Scratch: programming for all. Commun. ACM **52**(11), 60–67 (2009)
17. Upton, E., Halfacree, G.: Raspberry Pi User Guide. Wiley, New York (2014)
18. Vilner, T., Zur, E., Gal-Ezer, J.: Fundamental concepts of CS1: procedural vs. object oriented paradigm-a case study. ACM SIGCSE Bull. **39**(3), 171–175 (2007)
19. Weintrop, D., Wilensky, U.: To block or not to block, that is the question: students' perceptions of blocks-based programming. In: Proceedings of the 14th International Conference on Interaction Design and Children, pp. 199–208. ACM (2015)
20. Wilson, C.: Hour of code: we can solve the diversity problem in computer science. ACM Inroads **5**(4), 22–22 (2014)
21. Wing, J.M.: Computational thinking. Commun. ACM **49**(3), 33–35 (2006)
22. Wolber, D., Abelson, H., Spertus, E., Looney, L.: App Inventor. O'Reilly Media Inc., Sebastopol (2011)

Blexer-med: A Medical Web Platform for Administrating Full Play Therapeutic Exergames

Martina Eckert[1(✉)], Mónica Jiménez[1], María-Luisa Martín-Ruiz[2], Juan Meneses[1], and Luis Salgado[3]

[1] Centro de Investigación en Tecnologías Software y Sistemas Multimedia para la Sostenibilidad (CITSEM), Universidad Politécnica de Madrid, 28031 Madrid, Spain
martina.eckert@upm.es

[2] Departamento de Ingeniería Telemática y Electrónica, E.T.S. de Ingeniería y Sistemas de Telecomunicación, Universidad Politécnica de Madrid, 28031 Madrid, Spain

[3] Grupo de Tratamiento de Imágenes, E.T.S.I. de Telecomunicación, Universidad Politécnica de Madrid, 28040 Madrid, Spain

Abstract. The work presented here is part of the currently ongoing project "Blexer" (Blender Exergames), which aims at creating a complex exergaming environment for people with physical impairments. The games are thought for daily fitness exercises or rehabilitation and are based on movements captured with the Xbox 360 Kinect® sensor. Via a middleware, the games, currently implemented in the Blender Game Engine, communicate with a web platform called "Blexer-med", which is the working tool for the health professionals to manage the sessions of their patients. With help of the platform, clinicians can assign different games to their patients, personalize exercises by adjusting their difficulty, and obtain their performance results in real time. The advantage is that the exercises can be adapted by the therapist according to the results without the need to meet the patients, who do the exercise at their homes. This paper presents the architecture and functionalities of the web platform and its belonging database, as well as an example use case. The platform is currently tested by some volunteers suffering from muscular dystrophies, with a newly created adventure game. Although our work focuses on people with disabilities, the platform is generic and could be connected to any kind of exergames aimed at other purposes.

Keywords: Exergames · Web platform · Disability · Rehabilitation Kinect · Blender

1 Introduction

Physical Therapy and Fitness is important and necessary for everybody, as it provides opportunities for strength development, physical conditioning and functional training. Unfortunately, for people with physical impairments, the possibilities to do sports are

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 289–299, 2018.
https://doi.org/10.1007/978-3-319-76111-4_29

very limited, and even less if they are wheelchair dependent, therefore they usually are not able to do more than attending some days a week to physiotherapist sessions. To overcome this problem, exergaming, i.e. using active console video games that track player movements to control the game (e.g., Xbox-Kinect, Wii), may provide a form of exercise, if the games were configurable to the possibilities and needs of the disabled. Bonnechère et al. showed that video games, not especially designed for clinical purposes, are as efficient as conventional therapies or lead to even better therapeutic outcomes. They also have numerous advantages, such as preventing monotony and boredom, increasing motivation, providing direct feedback, and allowing double-task training [1].

There are many proposals in literature that present especially designed exergames for the rehabilitation of frequent diseases like Stroke or Parkinson's, which mainly affect elderly. To the best of our knowledge (we published an extensive state of the art section in [2]), there is no proposal for a general solution useful for a broad range of people with different disabilities, which, furthermore, goes beyond the possibilities of a mini-game, i.e. incorporates some kind of a history or purpose other than the exercise, to engage the player.

Regarding the remote configuration and management of the therapeutic games, which has the advantage that people can use them at their homes under supervision by the therapists, the solution is a web platform that can be connected with the gaming software. Some commercially available solutions have been found, e.g. MIRA Rehab Limited [3] and VirtualRehab [4]. Similar to the here presented work, the games register movements with help of the Microsoft Kinect camera and the clinician receives this data via the platform to evaluate the progress of the patient and gets the chance to reconfigure the adjustments. Both systems are applicable to different types of users like patients with Down syndrome, muscular diseases or cerebral palsy.

Furthermore, two similar systems have been developed in the same research group of Valladolid University: TELEKIN [5] and "La Isla EPIKa" [6]. The TELEKIN environment consists of four modules: an administrator web, a therapist web, a family web and the game module. The games work with either a Kinect sensor or a Leap Motion sensor to capture only the hand movements. "La Isla EPIKa" aims at the rehabilitation of stroke patients. Again, via a web platform, the therapist can adjust a number of difficulty parameters; the web offers also a guided mode for patients with severe paralysis.

Apart from those works, the European project REWIRE [7] (Rehabilitative Layout in Responsive Home Environment), funded by the seventh framework program, has to be mentioned. The consortium developed a rehabilitation platform for stroke patients based on virtual reality that allows the patients to do their exercises at home under remote monitoring by the hospital. The platform is composed of three parts: the patient station, the hospital station and the networking station. The patient station offers different simple exercise games, which are working with different sensors, as there is the Microsoft Kinect camera, a Balance Board and a haptic device. The clinician at the hospital who also receives the session output results, via the web, can configure them.

All mentioned systems have one thing in common: they manage a number of mini-games, i.e. games without interrelation as each represents a different exercise in a different environment and with an own purpose. The patients have to use those corresponding to their needs repeatedly to improve their physical state, but the exercise itself is not very different to doing it in a clinical centre together with a therapist. The difference of our work is to overcome this pure translation of exercises from the gym to a virtual environment. Instead, the exercises are integrated into a meaningful gaming environment, where the user is not aware of doing exercises but engaged to have the gaming experience and to win. This avoids that the patient gets bored and demotivated through too much repetitiveness and too evident exercises. Here, the application of gamification rules is very useful, as they help to create a gameplay that captures the players, such that they continue without being forced. As stated by Chou in his book "Actionable Gamification" [8], there are eight core drives that engage a person to do an activity, not only referring to games. A very important one is "Epic Meaning & Calling", which refers to the identification of the player with the main game character of a game and their feeling to be part of its history. As a consequence, the creation of full games with a complex history and meaning, seems to us a very promising way to really engage patients in doing their exercises, above all in long-time exercises, which are needed in case of chronical diseases.

Therefore, this work does not present yet another web platform. The main contribution of the here presented "Blexer-med" platform, which is part of the project "Blexer" (Blender Exergames), is to provide direct access to a personalized configuration of the gameplay of immersive large full play games. As opposed to mini-games, they rest upon multiple scenes and different levels, which camouflage the exercises as tasks to be performed by the game character. The patients, while playing, are not aware of doing rehabilitation, but are forced to do the exercises due to the gameplay, i.e. they are motivated to win and therefore have to do what the therapist adjusted for them via the web. In our future work, the interaction between the games and the web will be enhanced with intelligent algorithms to augment its utility for the therapist. The aim is to adapt the game flow according to the user's performance, mood and necessities in real time, keeping motivation always at a maximum level. Furthermore, it is planned that the game could act as an assistant for the therapist, by indicating possible reconfigurations of the game according to the user's performance.

In the following, the system architecture is outlined and a use case is presented to show the versatility of the platform.

2 "Blexer-med": Proposed System Architecture

Requirements

First, the "Blexer-med" web environment is a tool for clinicians that helps to configure and supervise the exercises of their patients like a remote control. It is not intended for the patients like in some of the formerly mentioned approaches, because the patients should only concentrate in playing and should not be aware of doing exercises.

For organizational reasons, three user roles have been defined in a hierarchical manner:

1. Super-Administrator (super-user with several functions):

- Create new game structures including: name, description, exercises, configurable parameters, etc. Furthermore, he/she is in close contact to the game developer for carrying out new configurations in the web environment based on the new game characteristics.
- Create Centers: For new patients and professionals to access to "Blexer-med".
- Register and unregister Center-Administrators, Clinicians and Patients.

2. Center-Administrator

- Register and unregister Clinicians and Patients.

3. Clinician (therapist or doctor)

- Register and unregister their patients.
- Configure the gaming parameters of their patients.
- Observe and comment the data of other patients of the same center.

Currently, there is one game in alpha phase connected to the web. It is an adventure game with a young amphibian called "Phiby" as the main character, who is exploring an unknown world and has to find the way out of a valley. Therefore, it has to overcome certain obstacles like lakes, rivers and steep slopes by diving, rowing a boat or climbing up a tree next to a slope. There are also sites to chop wood, which is gained and enables to build bridges over the rivers, or little huts to rest. When a hut is built, the game status can be stored like a level to continue there a next time. In-between the obstacles, some apples are placed to recover energy. Whenever one of the obstacles is passed, a new part of the map is discovered, so step by step the player gets aware of the surroundings and is motivated to keep on exploring the world and to bring "Phiby" to its destiny. A more detailed description of the game can be found in [9].

The movements necessary to do the tasks involved in the scenes have been formerly tested in form of independent mini-games by a group of volunteers with different disabilities who were mostly wheelchair dependent [2]. Results have proven that the movements are functional and safe, so they have been transformed to the here mentioned exercises.

The requirement for the web environment is to allow personalized configurations of the individual exercises of any game similar to the formerly described one. Therefore, a hierarchical structure is needed, consisting in a root, the depending exercises and a generic number of adjustable parameters:

Game A:

- *Exercise 1 (parameters 1 to 4)*
- ...
- *Exercise n (parameters 1 to 4)*

Here, each exercise would be a task to perform in the game, e.g. climbing. The parameters can be adjusted according to the difficulty level of the exercise, e.g. in case of climbing, the height of the tree and/or the time to perform it.

To allow the configuration of any type of game to any type of user, this structure has to be totally generic and extensible, such that it does not depend on the type or number of parameters handled in each exercise. In this way, the same game could be configured differently for any player.

Layered Architecture
Figure 1 shows the architecture of the currently implemented system environment. The patient's PC contains the Exergame, created in Blender [10] and a Middleware named "Chiro" [11] that is in charge to transmit the data obtained from different sensors which is used as input to control the games.

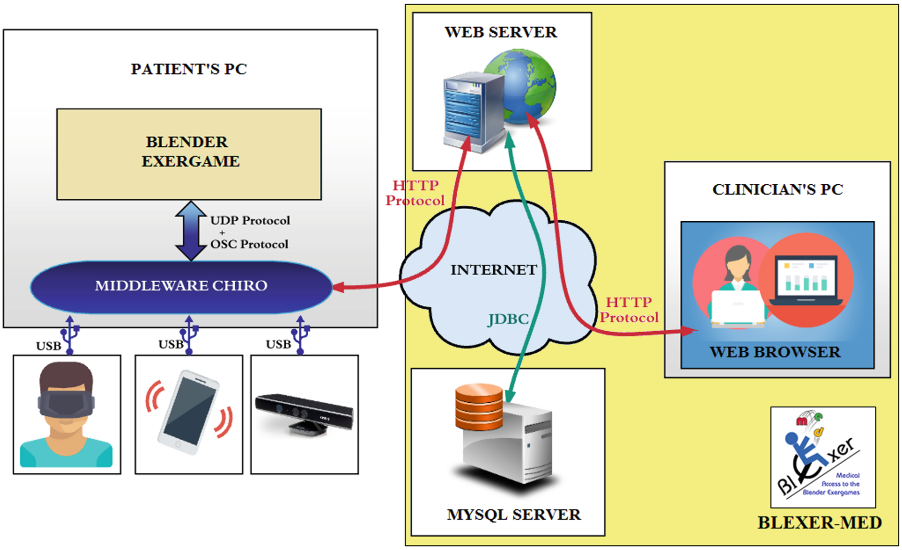


Fig. 1. "Blexer" system environment. On the left, the patients' PC represent the users at their homes; on the right, the "Blexer-med" system consists of a web server with the MYSQL database. The clinician's PC represent doctors and therapists at different medical centers.

With the aim of avoiding from the beginning the possible coupling between the different parts that compose the "Blexer-med" web environment, the web platform is structured into three layers:

(a) **A data layer** using MySQL server as an open source database manager. This layer stores all data belonging to the clinicians, administrators, center, patients, and game settings, etc. This layer provides information to the business layer in case of requests.
(b) **A business layer** implemented in Java using a RESTful web service (Representational State Transfer). The web service is accessed with help of the API JAX-RS (Java API for RESTful Web Services). This layer processes the user requests and accesses to the database.

(c) **A presentation layer** which is the heart of the distributed system that manages three types of interactions based on the client/server model. The aim of this layer is to maintain the system independent and easily adaptable to different platforms. The presentation layer shows the information to the user (clinician) and captures the events triggered by them.

As Fig. 1 shows these are the Protocols used in the different interactions:

1. **User** ← → **Web** interaction with HTTP (Hypertext Transfer Protocol) to access the platform
2. **Middleware** ← → **Web** interaction with HTTP to download game configuration and upload user results
3. **Database** ← → **Web** interaction with JDBC (Java Database Connectivity) to obtain the data to transmit to middleware or user.

Database

The database is a relational database and contains identification tables for administrators, centers, clinicians and patients on one hand, and game dependent tables for comments, game description, game setting, exercises, and session results. Two types of dependencies are used:

– **Identifying** (the "son's" existence depends on the "father's" existence), e.g. if a Center is deleted, also the administrator, clinicians and patients are deleted
– **Not identifying and independent** (a "son" would not be deleted in case the "father" was), e.g. Session Results are not deleted if a Game is deleted. This relation is needed to save performance data of the patients as long as they are registered on the platform, as it could be still of interest for the therapist.

Web representation

Regarding the presentation layer, Fig. 2 shows one example form, which is accessible by the Super-Administrator to define the exercises of the formerly mentioned adventure game. Some views of the clinician can be seen in Figs. 4 and 5 and are explained in Sect. 4.

3 Data Exchange for Game Configuration and Observation of Results

The configuration data set up by the clinician in the web environment, as well as the results obtained from the play, are stored in text files formatted in JSON (Java Script Object Notation) format to facilitate the exchange. As illustrated in Fig. 3, the middleware "Chiro" sends a HTTP GET request to the web service to download the configuration file, at the moment the user logs in.

If the user authentication was successful, the web service creates a JSON file containing the parameters marked by the therapist as "in use" (see also example in Fig. 4). The file contains a map for each game and inside a map for each exercise. A map is formed by a set of items, each consisting in the structure "key-value". The

Fig. 2. Super-Administrator view of the exercises defined for a game. Example for the game Phiby's Adventures which contains four exercises (Dive, Chop, Row and Climb). On the left, the tabs "Centers" and "Admin" give access to administrate health centers and their staff. "Games" gives access to administrate different games.

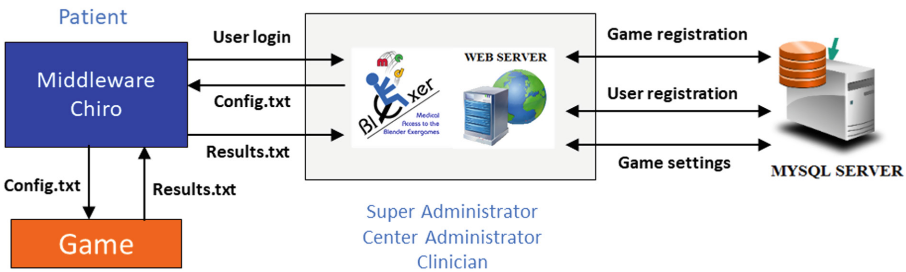


Fig. 3. Exchange of configuration information and results between the game and the web

key is an array of three strings: `[id_exercise, code, title]`; the value is obtained from the object created for the exercise. The file is then downloaded by "Chiro" and saved to the user's hard disc. Finally, the game itself de-serializes the JSON code and applies the values established for the different parameters in each exercise to the game flow.

For the interchange of the results obtained during the game play (scores, times etc.), the game creates a text file, also in JSON format, containing those values in form of an array. When the patient quits the game and closes “Chiro”, “Chiro” sends one result at each timer to the web platform, awaiting the confirmation of a correct de-serialization. In case of success, the result is deleted from the file and the next one will be sent. In this way, no results are lost due to bad internet connections. When a connection fails, the failed data will be re-sent the next time the user starts “Chiro”.

4 Example Use Case

In the following, we present a use case[1] to illustrate the particularities and advantages of our approach:

“Pepe Sanz” suffers a progressive muscular dystrophy and is using a wheelchair since 5 years. As a small boy, he could walk and play outside with his friends, now he is mostly alone and sitting in front of his computer. Unfortunately, the seated position is provoking a quicker degeneration of his muscle strength and a bad posture. His physiotherapist is visiting him twice a week and recommends daily arm and trunk movements to keep fit, but he is not motivated to do them, as they are very arduous and tiring. As he is a passionate gamer, he tried out the Xbox with Kinect and some sports and adventure games. But soon he got frustrated, as it was impossible for him to reach the same scores like his friends. One day, Maria, his therapist, came up with “Phiby’s Adventures”, connected it to his Kinect and calibrated the embedded movement amplifier to Pepe’s motion range. Through the web, she adjusted the parameters of the exercises as visualized in Fig. 4 for the exercise “Dive and eat” as an example. Figure 5 represents the view of the results for the same exercise.

In this exercise, “Phiby” has to be moved through the water with trunk movements to find a number of planktons to eat in a limit of time. Figure 6 shows a screen shot of the exercise.

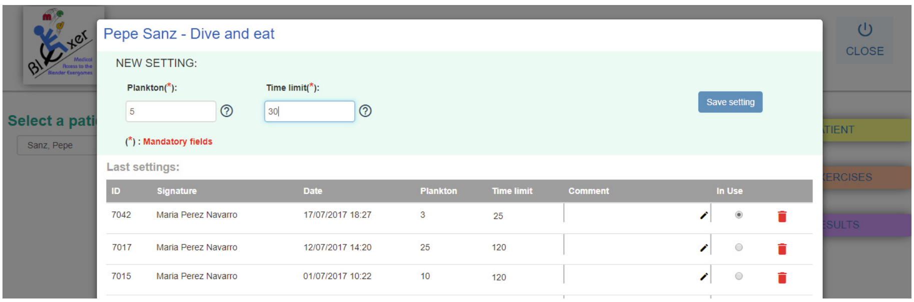


Fig. 4. Example for different configurations made over time for one exercise

---

[1] All names and personal data is invented and any possible relation to a real case is neither intended nor known. The performance data is also not real.

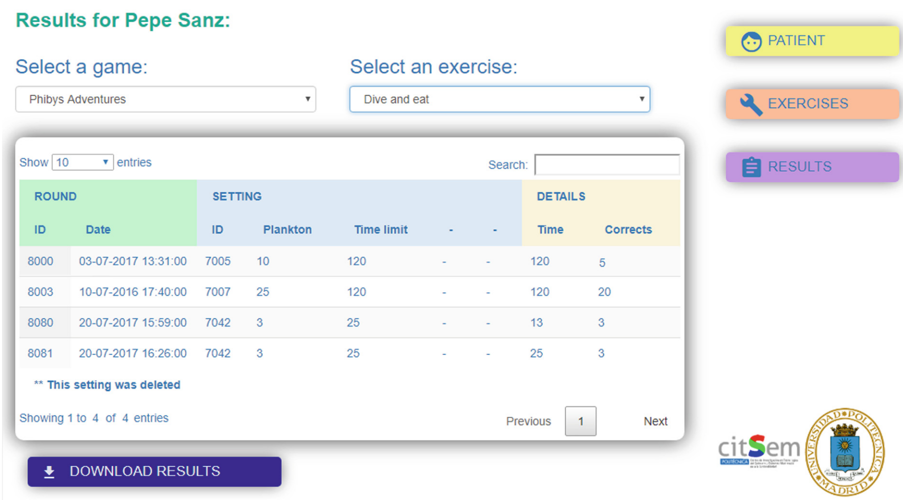


Fig. 5. Example for results obtained in different sessions performing the “Dive and eat” exercise. “Time limit” and “Time” in minutes.

Fig. 6. Screen shot of the “Dive and eat” exercise. The main character “Phiby” has to be moved with trunk movements (seated or standing) to catch the yellow pieces of plankton. (Color figure online)

It can be seen, that the initial setting of 10 pieces of plankton was too much, Pepe only got 5 and gave up after 20 min as it was the first time he played. Nevertheless, Maria augmented to 25 pieces and Pepe got 20, needing the maximum amount of time. As this was evidently too much effort, Maria set up only 3 pieces to get in 25 min. The first time, Pepe got them in 13 min, the next time, on the same day, he needed 25 min, because he got tired from the former play, where he also realized other exercises of the game. In short time he got engaged, because his therapist found quickly the right values for all exercises, and he played more each day and progressed such that his back muscles augmented in strength. This story is not real, and the values have been obtained while testing the environment, but we believe that they are useful to show how the tool could be used.

5 Conclusions and Future Work

In this work, we presented the web platform “Blexer-med”, which aims at assisting doctors and physiotherapists in the usage of exergames as an additional instrument for their therapies. The platform can be connected to customized games that incorporate different configurable exercises based on movements captured by the Kinect camera. Those games are not simple translations of real exercises to a virtual environment, but incorporate advanced game mechanics that enhance the user’s immersive feeling and motivation. Via the web-based platform “Blexer-med”, the therapist can adjust the difficulty of the exercises in relation to the patient’s needs and preferences, with the aim to maximize the user’s motivation to perform the exercises without being aware.

Currently, one full play adventure game that contains four different types of exercises, repeated in the order chosen by the patient, is tested together with the platform on a number of volunteers with disability suffering from rare diseases. Results will be hopefully published at the end of the year.

The next step of the project “Blexer” is the incorporation of an automatic adaptation of the game flow to the user’s performance and mood in real time. Therefore, the game will observe the user and adjust the parameters within the thresholds defined by the therapist, with the aim to keep the patient’s motivation at a maximum level. More advanced parameters regarding the exercises will be taken into account and decisions will be reported to the therapist via the platform.

Furthermore, a face recognition will be implemented to check if the right person is playing, a quality check of the movements and a more versatile configuration will be added and the processing of additional sensor data will be integrated to enhance the possibilities to check the user’s performance. All enhancement will then be tested with different types of user adaptive games.

Acknowledgements. This work was sponsored by the Spanish National Plan for Scientific and Technical Research and Innovation: TEC2013-48453-C2-2-R.

References

1. Bonnechère, B., Jansen, B., Omelinab, L., Van Sint Jan, S.: The use of commercial video games in rehabilitation: a systematic review. Int. J. Rehabil. Res. **39**(4), 277–290 (2016)
2. Eckert, M., Gómez-Martinho, I., Meneses, J., Martínez, J.F.: New approaches to exciting exergame-experiences for people with motor function impairments. Sensors **17**(2), 22 p. (2017). Article no. 354
3. MIRA Rehab. http://www.mirarehab.com. Accessed 28 June 2017
4. VirtualRehab. http://www.virtualrehab.info. Accessed 28 June 2017
5. TELEKIN. http://gti.tel.uva.es/juegos-serios-para-rehabilitacion-fisica-yo-cognitiva/telekin-sistema-de-telerehabilitacion-para-discapacidades-motoras-y-cognitivas. Accessed 02 July 2017
6. La Isla Epika. http://gti.tel.uva.es/juegos-serios-para-rehabilitacion-fisica-yo-cognitiva/la-isla-epika-juego-serio-de-realidad-virtual-para-rehabilitacion-de-ictus. Accessed 02 July 2017

7. REWIRE: FP7-ICT-2011 Call 7 - Personal Health System (PHS). https://sites.google.com/site/projectrewire/. Accessed 28 June 2017
8. Chou, Y.-K.: Actionable Gamificación. Beyond Points, Badges, and Leaderboards. Octalysis Media, Fremont (2015)
9. Eckert, M., Gómez-Martinho, I., Estéban, C., Peláez, Y., Meneses, J., Salgado, L.: Blexer – full play therapeutic blender exergames for people with physical impairments. In: Guidi, B., et al. (eds.) GOODTECHS 2017. LNICST, vol. 233, pp. 300–310 (2018). https://doi.org/10.1007/978-3-319-76111-4_z
10. blender.org. https://www.blender.org. Accessed 02 July 2017
11. Eckert, M., Gómez-Martinho, I., Meneses, J., Martínez, J.F.: A modular middleware approach for exergaming. In: IEEE International Conference on Consumer Electronics, Berlin, pp. 172–176 (2016)

Blexer – Full Play Therapeutic Blender Exergames for People with Physical Impairments

Martina Eckert[1(✉)], Ignacio Gomez-Martinho[1], Cristina Estéban[1], Yadira Peláez[1], Juan Meneses[1], and Luis Salgado[2]

[1] Centro de Investigación en Tecnologías Software y Sistemas Multimedia para la Sostenibilidad (CITSEM), Universidad Politécnica de Madrid, 28031 Madrid, Spain
martina.eckert@upm.es
[2] Grupo de Tratamiento de Imágenes, E.T.S.I. de Telecomunicación, Universidad Politécnica de Madrid, 28040 Madrid, Spain

Abstract. This work presents the "Blexer" (Blender Exergames) environment, which aims at the design and implementation of generic, adaptive and customizable rehabilitation exergames for people with physical disabilities. Currently, it consists in one full play exergame based on the movements captured by the Microsoft Xbox 360 Kinect camera and a medical platform that allows the remote configuration of the game. Opposed to multiple mini-games with specific purposes found in literature, here the focus is set on the design of a full game, which includes a character, a story and advanced game mechanics to achieve the integration of currently four exercises into a 3D adventure videogame. The principle aim is to enhance the patient's motivation to perform daily exercises with help of the game, therefore we apply the Core Drives defined in the Octalysis framework. Furthermore, the game incorporates a motion amplification functionality to augment the immersive feeling of disabled people in the video games. Last but not least, the configuration possibilities of the exercises included in the game are described.

Keywords: Exergames · Physiotherapy · Rehabilitation · Disability Kinect · Gamification · Octalysis · Adaptivity · Blender · Web platform

1 Introduction

For people with physical impairments, the possibilities to do sports are very limited. Nevertheless, they need to keep fit just as anyone else; even more, if they are wheelchair dependent because muscles tend to reduce with a lack of usage. Many people suffering from chronical diseases need to visit a physiotherapist multiple times a week, which is a costly and time-consuming burden.

Due to this situation, exergames are becoming more and more popular, as they offer a cheap possibility to exercise at home. Many scientific studies for application in rehabilitation systems have appeared recently, especially with the availability of 3D motion capture cameras like the Microsoft Kinect, proving that home exercising in

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 300–310, 2018.
https://doi.org/10.1007/978-3-319-76111-4_30

virtual reality environments is healthy and safe [1–3]. It has also been observed that video games, not specially designed for clinical purposes, are just as efficient as conventional therapies and can lead to even better therapeutic outcomes. Furthermore, games prevent monotony, increase motivation, provide direct feedback, and allow double-task training [2, 4].

Nevertheless, most proposals found in literature present exergames specifically designed for the rehabilitation of frequent diseases like Stroke or Parkinson's, which affect mainly the elderly. It is difficult to find solutions for a broader range of people with different disabilities. Furthermore, it is rare to find a solution that qualifies as more than a simple mini-game, i.e. incorporates some kind of story or purpose rather than the exercise itself, which would allow to engage the player and to motivate them to exercise without being conscious about it.

Engagement is a very important factor to increase the motivation of a person in many situations of life, as can be learned from Chou in his book "Actionable Gamification" [5]. Chou discovered the existence of eight main core drives (CDs) that usually engage people and analysed them in the so-called Octalysis framework:

CD 1 – Epic meaning & Calling
CD 2 – Development & accomplishment
CD 3 – Empowerment of Creativity & Feedback
CD 4 – Ownership & possession
CD 5 – Social Influence & Relatedness
CD 6 – Scarcity & Impatience
CD 7 – Unpredictability & Curiosity
CD 8 – Loss & Avoidance

They can be applied to every task to be gamified, so in this work, we took them as a guideline to create motivational factors in the game. CD 2 and CD 4 are the easiest and most straightforward elements to implement in a game: the goal to reach and the award obtained for it. However, the other core drives, although more subtle, are the ones that really engage the player. For example, CD 1 refers to a story and the identification of the player with a game character; CD 3 is expressed when users are engaged in a creative process; CD 7 keeps a player constantly engaged with the desire to find out what is happening next in the game.

So, the best way to really engage patients in their exercises, especially in long-time therapy (which is necessary in the case of chronical diseases), is to create full games with a complex story and purpose. In this sense, the first prototype of a full game, currently tested as an alpha version, has been created as part of the project "Blexer" (Blender Exergames), which aims at creating a complex exergaming environment for people with physical disabilities. The game is a follow up of our former work published in [6], where four mini-games have been tested with a group of people suffering from rare muscle diseases.

To the best of our knowledge, the work presented here is the first communication of a fully playable exercise game based on Kinect, as opposed to mini-games with very specific purposes. To give just two out of many examples: a rehabilitation game aimed at training dynamic postural control for people with Parkinson's disease is presented in

[7] and the authors of [8] designed “PhysioMate”, a game for stroke patients and elderly, that helps to train the upper-body for balance and motor coordination.

The definition of a mini-game can be found as: “A minor or incidental game within a larger video game.”[1] Furthermore, we observed that mini-games are simpler, less complex (usually only one scene, one character, one goal) and playable in a short time. Usually, there are no conditions that maintain the player engaged after finishing the game, other than beating their highest score. Nevertheless, in our opinion, exercises are too tiring for this and a better way would be to camouflage them in a large and compelling game environment. From now on, the authors will refer to a large game in contrast to a mini-game as a “full play game”. It could include the exercises in form of mini-games, but they could also be resolved implicitly (by moving a character, opening/closing menus etc.).

In the remaining sections, the complete “Blexer” system architecture will be presented. Afterwards, the design and structure of our first integrated full play therapeutic game are described. Furthermore, as a major concern is the flexibility of the games and their adaptability to any patient, taking into account their capabilities and needs, it will be focused also on two important functionalities:

1. Amplification of weak user movements to wider ones in the game character, with the aim to enhance the immersive feeling.
2. Configurability of exercise difficulty by a clinician remotely via a web platform.

2 “Blexer” System Architecture

The gaming environment presented here can be divided into two main parts as shown in Fig. 1: the user side (left), and the clinical side (right, yellow part) called “Blexer-med”, with multiple medical centers and as many therapists or doctors as needed.

The patient’s PC is working as the playing station, with the Exergame and a modular middleware called “Chiro” [9] that transmits the data obtained from one or more sensors to the game. The middleware could be configured for the Xbox 360 Kinect camera, a mobile phone and a VR headset [10] to capture, apart from the skeleton information provided by Kinect, additional movements like wrist and head rotation. In the future, it is planned to add more sensor devices like a heart rate measure and an oximeter, to obtain also information about the physical state of the patient with the aim to adapt the course of the game to their actual condition. Nevertheless, for the game presented here, we only use the Kinect data.

The right side of Fig. 1 shows how the medical access to the games is resolved in the “Blexer-med” web platform. It consists of a web server connected to a MySQL database which handles the communication between the clinicians and the middleware. The clinicians (therapists or doctors of different centers) can register their patients, assign one or more games of those available in the platform (currently only the

[1] https://www.wordnik.com/words/minigame

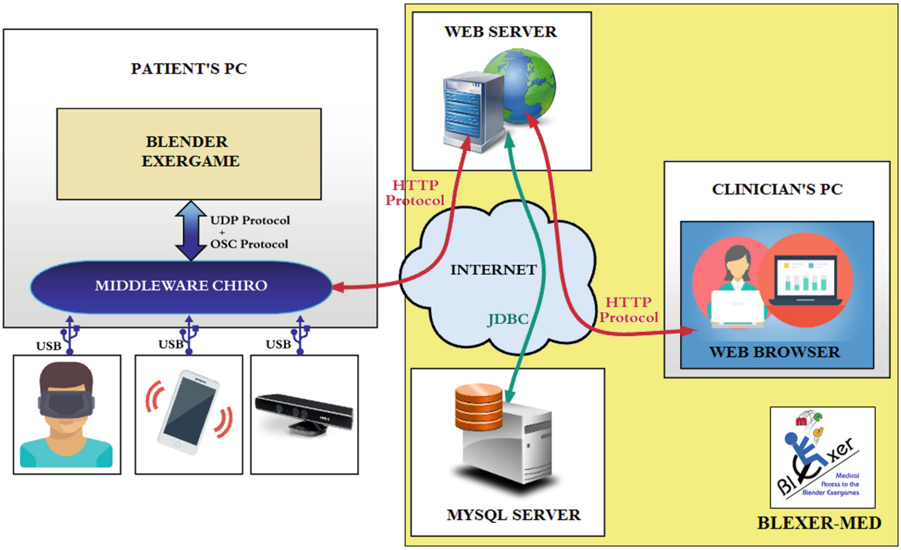


Fig. 1. "Blexer" system environment. On the left, the patients' PC represent the users at their homes; on the right, the "Blexer-med" system consists of a web server with the MYSQL database. The clinician's PC represent doctors and therapists at different medical centers. (Color figure online)

presented one) and adjust the parameters of the exercises involved in the games to the capabilities and needs of the patients.

When the patient starts the game at home, he/she first logs into the middleware, which then connects to the database and downloads the configuration information of the patient, stored in JSON (Java Script Object Notation) format, into a text file. After each playing session, the game writes the results (scores, playing times etc.) into a text file which is then uploaded by the middleware to the web platform. In case of a successful upload, the information is deleted from the file, if not, it is kept until the next successful internet connection. In this way, the clinician can analyze the results as soon as the patient terminated their exercise and evaluate their performance. According to it, the parameters can be readjusted for the next play.

The "Blexer" environment is modular and designed to manage an unlimited number of different games. Those games could be of any genre (action, adventure, role play, strategy etc.), as long as they are designed for a "corporal play", i.e. the player has to perform gestures and movements. Those inputs are modular and optional, and the prototype game presented in this article is only using the movements captured by the Kinect. It has been developed with the Blender modelling environment and game engine [11].

3 Full Play Therapeutic Exergame: "Phiby's Adventures"

3.1 Requirements

The goal of this work is the creation of a large game for people with disabilities, containing four different exercise types tested in [6], which have to be performed multiple times. The game should motivate and engage the user to maximize the playing time, therefore, the exercises have to be integrated into an interesting gameplay, which allows the repetition of the exercises without getting boring. To achieve this, specific game mechanics have to be added that motivate the player according to of the formerly introduced core drives.

3.2 Gameplay, Character, Story, World and Mechanics

The first decision to be taken before creating the game was choosing the right game genre. The decision fell on the adventure type, as it provides many possibilities to resolve the CD 1 (Epic meaning & Calling) with a cute game character and a nice story, as well as to integrate many tasks of different types, which give sense to the exercises.

As main character, a young amphibian called "Phiby" (see Fig. 2) has been created in different life stages (pollywog, toddler and infant), with the aim to represent the gaining of strength and skills due to the exercises (CD 2 – Development & accomplishment, CD 4 – Ownership & Possession). Nevertheless, in the game prototype presented here, it still does not appear in all forms.

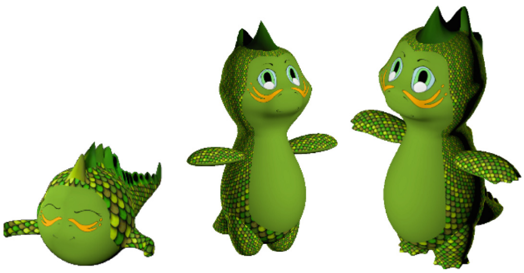


Fig. 2. "Phiby" in different states of development. Left: Pollywog. Center: Toddler. Right: Infant. In each state, his physique is developing a step further (growing limbs, fingers etc.).

The story can be resumed like: "Phiby" has fallen into an unknown world and has to find his family. In this first part, he explores the world around him to find the way out of a valley. To do so, he has to overcome certain obstacles like lakes ("Dive and eat"), rivers ("Row the boat"), and slopes ("Climb the tree"). There are also fallen trees that can be chopped to obtain pieces of wood ("Chop the wood"), which later can be used to build bridges over the rivers or little huts to rest and save the progress. These tasks are the physical exercises to be performed by the player (exercise names in brackets) and described in Sect. 3.3 in detail.

To achieve a repetitive execution of the exercises without getting bored, a game flow has been designed that lets the user decide which tasks to do next, therefore he/she will be motivated by the following additionally integrated game mechanics. Most of them are objectives and awards that provoke CD 4 (Ownership & possession) and CD 6 (Scarcity & Impatience), but some also tackle others as indicated:

- **Wood.** Every chopping exercise rewards 7 kg of wood.
- **Bridges.** 5 kg are required to build a bridge over a river, such that there is no need to row over it. A bridge allows also to go back to the former cell.
- **Huts.** Every 10 kg of wood allow building a hut in a specific point of the map. At this point, the status of the game will be saved, such that the user can resume the game from that point on.
- **Map.** Whenever one of the tests is passed, a new area of the "world" is discovered, so step by step the player gets aware of the surroundings and is motivated to keep on exploring the landscape and to bring "Phiby" to his destination (CD 1 – Epic meaning & Calling). Whenever a hut is built, a larger part of the map is discovered, such that it is easier to find the way to the end (CD 7 – Unpredictability & Curiosity).
- **Apples.** In-between the obstacles, some apples are placed to recover the energy lost during the exercises. Every apple recovers 10% of the energy lost in an exercise.

Figure 3 shows a simplified diagram of the game structure as it is currently implemented. The layout of the current game environment is stored in an XML file and consists of a map of 16 × 8 square cells, each representing a patch of the landscape to be explored. The XML parser reads the information of the scenes to be built from the file and passes it to the game engine. Every cell has one obstacle (lake, trunk, river, tree or a wall) in each border, which has to be passed to enter the adjacent cell by executing the corresponding exercise, DIVE, CHOP, ROW, and CLIMB, except for the wall,

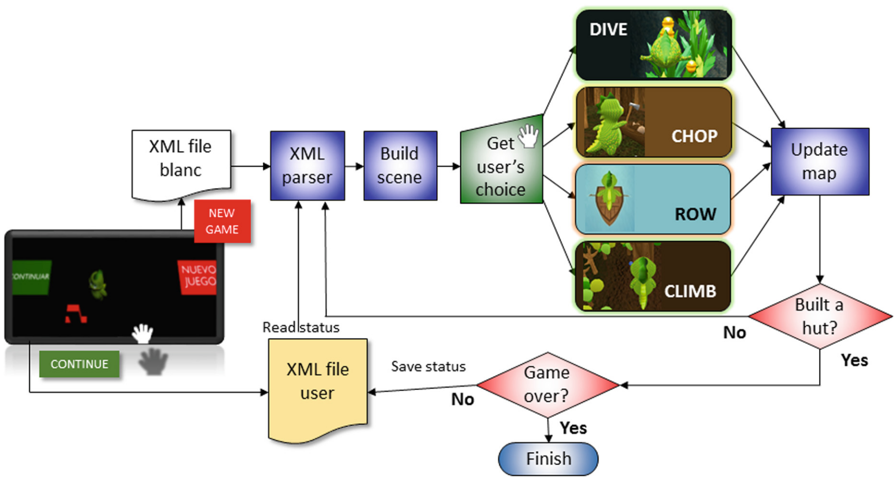


Fig. 3. Game outline of "Phiby's Adventures". (Color figure online)

which is an insurmountable barrier. The user can freely choose where to go by moving a red arrow with the right hand, which is followed by "Phiby". When "Phiby" reaches the arrow close to an obstacle, the scene of the corresponding exercise is loaded.

Figure 4 shows the scenes built out of two cells: the left one with a river, a lake, a wall and a tree, the right one with a trunk, a lake, a wall, and already chopped wood to the south. "Phiby" is following the red arrow, moved by the user's right hand. Once reached the arrow placed on an obstacle, the corresponding exercise starts. The right image also shows the map of the already explored cells, which can be opened by raising the left hand. The game is then paused until the map is closed by raising the other hand. The map shows the explored parts. Next, "Phiby" can go to the west to chop more wood or to the east to dive into the lake. The title of Fig. 4 explains the information presented to the player.

Fig. 4. Screen shots of intermediate scenes. Left: a river (Row the boat), a tree (Climb the tree), a lake (Dive and eat), and a wall (no passing). The energy is on top and still there has no wood been gained to build a bridge or a hut. Right: "Phiby" just passed the "Chop the wood" exercise and gained 7 kg. The energy is at 90%, but there is an apple to eat to increase the energy. (Color figure online)

3.3 Technical Description of the Exercises

As stated at the beginning, apart from enhancing the motivation of the users to play and do their exercises, two more objectives are inherent to the work presented here:

(1) Augment the immersive feeling during the play by amplification of the user's movements.
(2) Adapt the difficulty of the exercises to the user's possibilities and needs by adjustment of the corresponding parameters.

The amplification of the user's movements means that the real movements are represented as larger ones by the game character as if they were simply copied. If the game character has to raise an arm to get an apple, but the patient is not able to raise it as high as needed, the maximum point of the patient's arm movement is calibrated to the maximum point of the game character's arm movement. In this way, every movement is scaled inside the game and appear the same as if a user without physical restrictions were playing. Figure 5 visualizes this principle and how it is applied to the skeleton of “Phiby”. To achieve these scaling factors, a first step is always the measurement of the user’s maximum motion ranges of all limbs, head and trunk [6, 9].

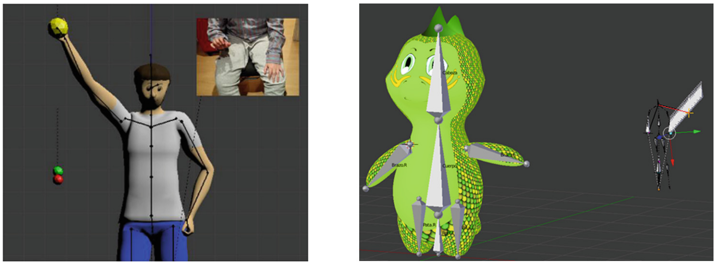


Fig. 5. Left: illustration of motion amplification principle: the user’s real maximum movement (photo) corresponds to the green sphere of the avatar, the yellow sphere marks the scaled movement. Right: visualization of the character’s skeleton and in the background the (invisible) Kinect skeleton with amplified motion transferred to the arm. (Color figure online)

The four exercises implemented in the game and the parameters needed to adjust their difficulty are described in the following subsections. Figure 6 illustrates some screen shots taken during the play.

Fig. 6. Scene shots of the currently implemented exercises: (a) “Chop the wood” (3 of 8 trunks chopped) (b) “Dive and eat” (2 of 8 planktons eaten) (c) “Row the boat” (distance left to the shore: 12 m) and (d) “Climb the tree” (7 m left to the top)

Chop the wood

The movement to be performed in this exercise is an up and down movement of the right arm. The user has to keep the arm up during a certain time to "charge" the axe which is visualized with a silver ring around the axe as shown in Fig. 7 on the left. When it is charged, a flash appears and the user can move the arm down to chop the wood (Fig. 7, 2nd image). The focus of the exercise is set on endurance, precision and speed are not necessary. It's more, patients with muscular dystrophies should try to not simply let the arm fall down. Therefore, the difficulty of this exercise can be adjusted with four parameters: height to raise the arm, duration keeping the arm up, time to hit the trunk, and the number of pieces to chop.

Fig. 7. Details of the scenes. Left: the axe in "Chop the wood" when it is charging and when it is ready to use. Right: a piece of plankton in the "Dive and eat" exercise. (Color figure online)

Dive and eat

The diving exercise focuses only on the movement of the trunk, in the directions forward, backwards and sideways. This is a very important exercise for wheelchair users, as they have to strengthen their back and pelvic muscles. In this exercise, the pollywog is diving through a lake and its direction is controlled by the user's movements. The goal is to capture the pieces of plankton floating in the water (Fig. 7, right image) and to eat them. The unique parameter to adjust is the number of planktons to control the duration of the exercise. It could be considered to implement a speed factor to enhance the movements in case the therapists considered this to be useful.

Row the boat

In this scene, the character is sitting in a rowing boat and has to cross a river. Here, the only adjustable parameter is the distance to the shore. To push the boat, both arms have to be moved back and forth, without the need of an exactly symmetrical movement. Currently, the boat is moving quicker with a quicker movement, so the user is engaged to finish the exercise earlier. As further possible enhancements of difficulty, obstacles or wild animals could be added in the water to be avoided or escaped.

Climb the tree

Here, "Phiby" has to climb up a tree, which is adjustable in height. In the current environment, the goal is to overcome a wall next to the tree; but the movement is adaptable to other scenes, i.e. to climb up a steep face or to get out of a hole. The only adjustable difficulty parameter is the height of the tree, which implicitly determines the number of arm movements to be made. The movements have to be alternated, where a blue ball appears to indicate the hand the user needs to raise at the moment. When a movement is successful, the character is animated and climbs a step up. Another possible parameter to implement would be the speed of the movements, i.e. the time it takes to reach the top.

4 Conclusions and Future Work

In this work, we present the "Blexer" system together with the first prototype of a full play game for therapeutic purposes for persons with physical impairments. The system described here is currently tested in its alpha phase to reveal and fix consistency errors. Afterwards, long term tests will be driven with a greater number of volunteers. Results will be analyzed and compared with former results.

In our ongoing work, we are currently porting the system to the Unity 3D game engine and also integrating Xbox One Kinect v2 body and face data. The middleware will be enhanced with other sensors as heart frequency measure and oximeter to capture information about the user's physical state. Together with an emotion detection process, the possibilities to realize a Dynamic Difficulty Adjustment (DDA), i.e. to adapt the pace and course of the game intelligently to the user's performance and mood, will be studied to achieve a maximized motivation.

Further on, the story will be developed with more detail, including additional NPC (non-player characters), further stages of development and advanced exercises. Also, the possibility to add a multiplayer mode will be considered, as this would potentiate a further core drive as a motivational factor: CD 5 (Social Influence & Relatedness).

Acknowledgements. This work was sponsored by the Spanish National Plan for Scientific and Technical Research and Innovation: TEC2013-48453-C2-2-R.

References

1. Webster, D., Celik, O.: Systematic review of Kinect applications in elderly care and stroke rehabilitation. J. Neuroeng. Rehabil. **11**(108), 1–24 (2014)
2. Pompeu, J.E., et al.: Feasibility, safety and outcomes of playing Kinect Adventures!TM for people with Parkinson's disease: a pilot study. Physiotherapy **100**, 162–168 (2014)
3. Hondori, H.M., Khademi, M.A.: Review on technical and clinical impact of Microsoft Kinect on physical therapy and rehabilitation. J. Med. Eng. **2014**, 846514–846530 (2014)
4. Bonnechère, B., Jansen, B., Omelinab, L., Van Sint Jan, S.: The use of commercial video games in rehabilitation: a systematic review. Int. J. Rehabil. Res. **39**(4), 277–290 (2016)
5. Chou, Y.-K.: Actionable Gamificación. Beyond Points, Badges, and Leaderboards. Octalysis Media, Fremont (2015)
6. Eckert, M., Gómez-Martinho, I., Meneses, J., Martínez, J.F.: New approaches to exciting exergame-experiences for people with motor function impairments. Sensors **17**(2), 22 p. (2017). Article no. 354
7. Madeira, R.N., Costa, L., Postolache, O.: PhysioMate-pervasive physical rehabilitation based on NUI and gamification. In: Proceedings of the International Conference Exposition Electrical Power Engineering, Iasi, Romania, pp. 612–616 (2014)

8. Galna, B., et al.: Retraining function in people with Parkinson's disease using the Microsoft Kinect: game design and pilot testing. J. Neuroeng. Rehabil. **11**, 60 (2014)
9. Eckert, M., Gómez-Martinho, I., Meneses, J., Martínez, J.F.: A modular middleware approach for exergaming. In: IEEE International Conference on Consumer Electronics, Berlin, pp. 172–176 (2016)
10. Eckert, M., Zarco, J., Meneses, J., Martínez, J.-F.: Usage of VR headsets for rehabilitation exergames. In: Rojas, I., Ortuño, F. (eds.) IWBBIO 2017. LNCS, vol. 10209, pp. 434–442. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56154-7_39
11. https://www.blender.org. Accessed 02 July 2017

Smart City and Images: The Use of Image Hashtags to Get Insights on Citizens

Roberta De Michele[1] and Marco Furini[2(✉)]

[1] Fondazione Marco Biagi, Università di Modena and Reggio Emilia, 42121 Modena, Italy
roberta.demichele@unimore.it
[2] Dipartimento di Comunicazione ed Economia, Università di Modena and Reggio Emilia, 42121 Reggio Emilia, Italy
marco.furini@unimore.it

Abstract. The knowledge of citizens' interests and problems is crucial for the smart management of a city. Up to few years ago, the employment of opinion agencies was mandatory to get this knowledge, but since this process is cost and time consuming, many studies are beginning to exploit social media contents to understand citizens. In particular, the attention is usually focused on textual data, and only few studies consider multimedia contents. In this paper, we investigate whether it is possible to know citizens' interests and problems by using images published in the Instagram platform. In particular, we propose a method that analyzes and measures the importance of hashtags associated to images. The experimental evaluation shows that images could be an important source of information to understand citizens.

Keywords: Smart city · Image analysis · Hashtag classification Hashtag indexing

1 Introduction

The Internet of Things (IoT) is expected to change the way we live and work. It is a paradigm that envisions a future where people and objects will be interconnected and will be able to communicate each other through suitable protocol stacks [1,2]. Currently, there are millions of physical objects already connected to the Internet and different research reports estimate that from now on millions of new objects will enter the IoT world every day, reaching the number of 21 billion by 2020 and generating a market of more than $300 billion by 2020. These objects will contribute to make the Internet even more pervasive than it is today, and they will foster the development of applications able to provide new services to citizens, companies, and public administrations [3–5].

The transformation of a city into a smart city seems to be the solution to the problems emerging from rapid urbanization and urban population growth, as the employment of smart strategies might reduce wastage (e.g., energy, water,

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 311–320, 2018.
https://doi.org/10.1007/978-3-319-76111-4_31

etc.) and might improve services (e.g., public and private transportation, waste management, etc.), resulting in an improvement of the life quality of citizens [6]. Most of the processes, which aim to transform cities, primarily focus on the technological aspect (e.g., sensors, devices, communication protocols, etc.) but, although this is necessary, it is not sufficient to transform the city as a whole. Indeed, to really improve the services offered to citizens and to address challenges of development and sustainability more efficiently, citizens should be involved when designing and implementing strategies to transform a city into its smart version [7–9].

The involvement of citizens necessarily passes through their understanding and usually this task is usually achieved through the employment of opinion agencies that prepare and submit questionnaires to a sample of citizens. In general, these investigations are effective and regard general aspects of daily life (e.g., public transport satisfaction, social safety, etc.); however, they are expensive and time consuming. Therefore, many administrators tend to use these investigations sparingly. Recently, to overcome these problems, researchers are proposing an alternative approach: exploit the large amount of data available in social media platforms to understand citizens [10]. The goal is to use contents published in social media to make decisions that will end in better use of resources, better organization, better citizen lifestyle, better human relations and, eventually, better society.

In literature, one of the most used approach to understand citizens is the analysis of their sentiment in general, or towards specific topics [11–14]. For instance, Nakov et al. [15] exploit tweets and SMS contents to provide insights about the sentiment of citizens; Mitchell et al. [16] used geolocated tweets to measure the social level of happiness in specific locations; Lin [17] used Twitter data to identify the sentiments of people living in the Pittsburgh area; Guo et al. [18] used tweets to understand the influence of socioeconomic and urban geography parameters on happiness. Due to the large usage of multimedia contents [19,20], recently, researchers are considering the analysis of images posted by users on social platforms [21,22].

We are convinced that images can be transformed into an important source of information and, therefore, we propose a method that focuses on the hashtags present within the captions of images and has the goal to identify citizens' problems, topics and places of interests. It is worth recalling that hashtags are community-driven keywords used by people to emphasize contents and to highlight topics; originally introduced in Twitter, hashtags are spreading throughout the entire social scenario. Motivated by the large use of mobile photo sharing applications [23,24], in this paper we propose ICA (Image Caption Analysis), a method that measures and ranks the importance of hashtags within a set of images. By considering the hashtags related to problems, topics and places of interest it is possible to understand what citizens think. To evaluate our proposal, we performed an experimental assessment focused on the Instagram platform. We observed more than 110,000 images related to different Italian big/medium cities (e.g., Milan, Bologna, Florence, Rome and Naples) for a period of 20 days.

The obtained results show the feasibility of our proposal and also show that it is possible to identify citizens' problems and interests through an analysis of the captions of the images. Therefore, the proposed approach might be very useful to city administrators and/or enterprises that want to manage/transform a city into a smart city.

The paper is organized as follows: in Sect. 2 we briefly overview studies that use social media data to understand citizens; in Sect. 3 we present details of our ICA method, which is evaluated in Sect. 4. Conclusions are drawn in Sect. 5.

2 Related Work

Recently, many studies focused on social media data to measure the emotion of people who share data on public platforms. This information is assuming more and more importance in the nowadays scenario, as it might be used to provide better services and/or to manage resources in an urban context. For instance, through this "sentiment analysis", administrators can become aware of unhappy areas, might analyze and identify problems and might implement improvements in services such as public transport, security, recreation, etc. While there are several studies that focused on Twitter, the use of images as a source of information to retrieve insights about citizens is in its early stage and in the following we provide a brief overview of recent studies in this area.

Different proposals focused on images to retrieve information that can be used in the tourist sector. For instance, Xu [25] proposed a method that uses the photos users share in image-sharing websites (e.g., Flickr and Panoramio) to recommend tourist locations. The method does not analyze the contents of the image, but it focuses on the associated metadata: the embedded GPS data are used as the building blocks for the recommendation system. Bojic et al. [26] considered Flickr as a proxy for attractiveness. In particular, they investigated how country attractiveness scales with its population and size using the number of foreign users that take photographs. The proposed method did not analyzed images, but it extracted GPS data to infer the nationality of photographs and to compute the country attractiveness index. Quercia et al. [27] focused on the Flickr platform to design a method that automatically suggests routes that are not only short but also emotionally pleasant. Although based on an image-sharing platform, the proposed method did not consider image contents, but instead they exploited Flickr's metadata. You et al. [28] proposed a methodology to numerically represent the happiness of a city by mining user generated contents in Flickr. In particular, the methodology is based on titles, tags, descriptions, and comments. The obtained results are then compared against real world phenomena including population, crime rate, and climate. Bujari et al. [29,30] proposed to automatically modify the story of a trip by identifying interesting places to visit in the nearby of an itinerary. The developed system combines data associated to Flickr geo-tagged pictures with information provided by WikiLocation. A user study showed that users deeply appreciated the possibility of discovering interesting places located in the nearby of a given itinerary.

Other studies proposed a different focus: the usage of images to retrieve information that can be helpful to city managers. For instance, Abdullah et al. [31] analyzed images available on the Twitter platform, used them as a resource for sensing societal characteristics, and introduced a formalized measure of societal happiness. In particular, they converted to grayscale all the retrieved images and then they applied a smile detection algorithm to detect the happiness of people over a period of time. Liu et al. [32] proposed a method to filter redundant information produced by citizens when using a collaborative system provided by city administrators. They grouped together similar images by analyzing their RGB color histograms and then they extracted all the captions associated to similar images, with the assumption that similar images discuss similar topics. Oliveira and Painho [33] proposed to extract data from Twitter, Flickr, Instagram and Facebook in order to identify the Lisbon citizen's feelings and to help the smart city transformation process by improving urban functions and services. Unfortunately, there are no details regarding the mapping between images and citizens' emotions.

3 Image Caption Analysis

Motivated by the need to understand citizens through social media data and based on the observation that images are more and more used in the social media scenario, in this paper we propose ICA (Image Caption Analysis), a method that focuses on the hashtags associated to images to get insights about citizens.

Our hypothesis is that by collecting a set of images in a specific area (e.g., a city, a district, a street, etc.), the hashtags associated to the images represent an important source of information to identify citizens' problems, topics and places of interests, information that can be very useful in decision making processes. Indeed, as shown in Fig. 1, ICA takes in input a set of images, analyzes all the hashtags available in the set of images and ranks them according to their importance; then it outputs three different lists of information about citizens.

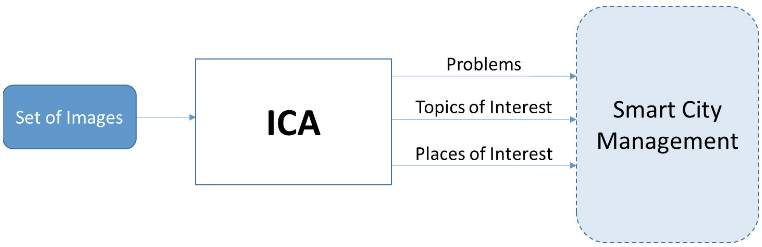


Fig. 1. ICA Scenario: ICA can transform a set of images into useful information for smart city administrators.

The core of the ICA proposal is the measurement of the importance of a hashtag. ICA does not use the number of times a specific hashtag appears within the dataset to define its importance, as this approach may be misleading. For instance, one may think of a single user who post several images with the same hashtag (e.g., `#smog`). Although this is the concern of a single user, the analysis would reveal that "smog" is a concern for the whole society. Similarly, suppose the dataset is composed of just two images: one described with the hashtag `#food` and one described with the hashtag `#art`. If we consider the number of times the hashtags appear within the dataset, it is impossible to call for the preferred topic.

A possible solution to deepen the understanding of the scenario is to consider the number of likes an image receives. For instance, suppose the `#food` image received 100 likes and suppose the `#art` image received no likes, is it still impossible to call for the citizens preferred topic?

In the analysis of hashtags, ICA considers the number of likes a picture receives and defines $Importance(H)$ as a way to measure the significance of a hashtag H within a set of N images:

$$Importance(H) = \sum_{i=1}^{N} P_i * L_i \qquad (1)$$

where $P_i = 1$ if H is a hashtag associated to image i, 0 otherwise; L_i is the number of likes that image i received.

By analyzing and ranking all the hashtags present within the considered dataset, ICA is able to measure the importance of these hashtags. Furthermore, by providing ICA with a set of possible problems (i.e., smog, security, pollution, decay, traffic, etc.), topics (i.e., art, food, architecture, museum, gallery, etc.) and places, ICA can output different lists of information about citizens' problems, topics and places of interests. Therefore, if a city administrator is interested in the opinions of citizens living in a specific area (e.g., a particular district, a block, a neighborhood, etc.) it is sufficient to collect the images produced in that area and to analyze them through ICA.

4 Experimental Analysis

The first mandatory step to investigate the feasibility of our approach, is data collection. We focused on the images related to big and medium Italian cities (i.e., Milan, Bologna, Florence, Rome and Naples) posted on the Instagram platform. In a period of 20 consecutive days, we observed 110,572 images. Details of the dataset are reported in Table 1, where we also show the number of hashtags used to describe these images.

Looking at the dataset, it is possible to note the large number of hashtags (i.e., on average from 21 to 30 per image) used to describe every image. This confirms that Instagram users make a great use of hashtags to enrich and complete the visual content of images. Therefore, hashtags represent an important source of information to get users' insights.

ICA extracts all the hashtags used to describe the images of the dataset, and it computes the $Iimportance$ index defined in Eq. (1). The resulting list of

Table 1. Dataset details.

City	Number of images	Avg number of hashtags per picture	Number of different hashtags
Milan	28,761	23	66,927
Bologna	13,114	21	33,376
Florence	17,718	21	37,507
Rome	29,865	24	68,758
Naples	21,114	24	47,918

hashtags is then sorted according to the *Importance* index (from the most to the less important one). Then, the hashtags related to problems, topics and places of interests are extracted. It is worth noting that the extraction is currently based on a predefined set of topics and of problems, whereas the extraction of places of interests is based on a separate file that is manually edited using the "Things to Do" list available in TripAdvisor.com. As a future work we plan to automate this process.

Table 2 summarizes the obtained results for the three different categories. Note that the number in parentheses indicates the ranking position, and also note that the ranking position is city-based and not category-based (i.e., ICA produced a list for every analyzed city).

Table 2. Ranking of the identified feelings, problems, topics and places of interests. The number in parentheses indicates the ranking position.

City	Topic of interest	Place of interest	Problem
Milan	fashion (6)	duomo (58)	traffic (4486)
	football (26)	galleriavittorioemanuele (97)	trash (8179)
	design (29)	navigli (329)	smog (10408)
Bologna	architecture (41)	piazzamaggiore (117)	trash (5288)
	fashion (45)	sanLuca (526)	traffic (6113)
	music (50)	sanPetronio (698)	smog (9608)
Florence	architecture (14)	duomo (22)	security (11472)
	art (20)	pontevecchio (41)	smog (17785)
	fashion (53)	arno (213)	trash (22603)
Rome	art (43)	coliseum (94)	traffic (3045)
	fashion (44)	vatican (137)	security (10451)
	architecture (45)	fontanaditrevi (241)	squalor (17089)
Naples	football (10)	teatrosancarlo (501)	fine (89)
	fashion (18)	lungomare (525)	traffic (8470)
	food (28)	casteldellovo (619)	trash (10112)

It is to note that we did not perform any tag filtering and therefore the ranking list also has hashtags related to applications used to modify the image (e.g., `#instagood`, `#instagram`), hashtags related to brand (e.g., `#nikon`, `#canon`, etc.), hashtags associated to specific events or to specific entertainment contents (e.g., `#xfactor`, `#bolognainspired`, etc.) and general hashtags (e.g., `#Italy`, `#girl`, `#city`, etc.).

A general look at the results shows that users are not really concerned about problems. In general, problems are very low in the ranking. For instance, in Milan the first problem (traffic) is in position 4.486, in Florence the first problem is security, but it is in position 11.472. The same applies to Bologna (trash in position 5.288) and Rome (traffic in position 3.045). The only exception is about the city of Naples: citizens pay particular attention to fines (position 89).

Results related to the topics of interest show what really matters to citizens. For instance, in Milan "fashion" is high ranked, whereas in Naples people are more interested in "football". "Architecture" is a popular interest both in Florence (14) and in Bologna (41). Citizens of Naples are also very interested in "food", whereas Bologna citizens are interested in "music".

Results related to the places of interest show the most popular places of the city. In addition to the places listed in the ranking, it is interesting to look at their position. For instance, in Naples, the "San Carlo" theater is the most popular venue, but it is ranked at position 501. Conversely, in Florence, the "duomo" (cathedral) is in position 22. It is also interesting to notice the difference within the same city. For example, in Bologna, "Piazza Maggiore" is the highest ranked venue (position 117), but the near "piazza Santo Stefano" (fifth ranked venues, not reported in the Table) is ranked 1029. The two places are 200 meters away, and clearly this indicates a problem in the communication process towards tourists: tourists (and people in general) seem to ignore this place.

In summary, results obtained through the experimental analysis confirmed that a smart analysis of the hashtags provides useful insights about citizens' problems, topics and places of interest. Although it is difficult to show the importance of the results, as these results are very linked to the knowledge of the city, it is possible to identify two large areas that could benefit from the results produced by the proposed ICA.

– **Smart City**. City Administrators can better manage resources and services as the knowledge of problems, topics and places of interest might be very helpful in decision making processes. For instance, they can monitor the city in a very precise way and they can take effective countermeasures (e.g., more trash-bins where there is a garbage problem, installation of surveillance cameras where citizens perceive a security issue, etc.). To make a particular example, by looking at the places of interest, Bologna administrators can act on public transportation to provide novel services and/or to improve the existing ones (e.g., "San Luca" is a church ranked second among the places of interest, but it is a venue not easy to reach as it is located outside downtown on atop a forested hill). Similarly, Bologna administrators needs to investigate the reasons why two famous venues (i.e., "San Petronio" and the "Towers") are not high ranked. In Florence, administrators needs to improve tourist routes (e.g., "Palazzo Vecchio" is ranked to 221 although very close to "Duomo" that is ranked 22).

- **Business**. Companies may take advantage of the knowledge of topics and places of interest to improve or create business opportunities. For instance, local businesses like restaurants and bars may find interesting business prospects around places with the highest ranking; similarly, new business opportunities can be developed around topics in the highest ranking. To make a particular example, in Naples people are interested in "food", whereas in Bologna there might be business opportunities in the area of music.

5 Conclusions

In this paper, we investigated whether it is possible to use images posted on the Instagram platform to know the interests and problems of citizens. We considered the hashtags present in the captions of the images, and we proposed a methodology to measure the importance of hashtag within a dataset of images. Results revealed the feasibility of the proposed method and confirmed our initial hypothesis that images could be an important source of information for city administrators and/or enterprises.

References

1. Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M.: Internet of Things for smart cities. IEEE Internet Things J. **1**(1), 22–32 (2014)
2. Atzori, L., Iera, A., Morabito, G.: The Internet of Things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010)
3. Furini, M., Mandreoli, F., Martoglia, R., Montangero, M.: IoT: science fiction or real revolution? In: Gaggi, O., Manzoni, P., Palazzi, C., Bujari, A., Marquez-Barja, J.M. (eds.) GOODTECHS 2016. LNICST, vol. 195, pp. 96–105. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61949-1_11
4. Mirri, S., Prandi, C., Salomoni, P., Callegati, F., Melis, A., Prandini, M.: A service-oriented approach to crowdsensing for accessible smart mobility scenarios. Mob. Inf. Syst. **2016** (2016)
5. Bujari, A., Furini, M., Mandreoli, F., Martoglia, R., Montangero, M., Ronzani, D.: Standards, security and business models: key challenges for the IoT scenario. Mob. Netw. Appl. **22**, 1–8 (2017)
6. Mohanty, S.P., Choppali, U., Kougianos, E.: Everything you wanted to know about smart cities: the Internet of Things is the backbone. IEEE Consum. Electron. Mag. **5**(3), 60–70 (2016)
7. Amaba, B.A.: Industrial and business systems for smart cities. In: Proceedings of EMASC, EMASC 2014, pp. 21–22. ACM, New York (2014)
8. Komninos, N., Tsarchopoulos, P., Kakderi, C.: New services design for smart cities: a planning roadmap for user-driven innovation. In: Proceedings of WiMobCity, WiMobCity 2014, pp. 29–38. ACM, New York (2014)
9. Roccetti, M., Marfia, G., Palazzi, C.E.: Entertainment beyond divertissment: using computer games for city road accessibility. Comput. Entertain. **9**(2), 10:1–10:9 (2011)

10. Bujari, A., Furini, M., Laina, N.: On using cashtags to predict companies stock trends. In: 2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC), pp. 25–28, January 2017
11. Montangero, M., Furini, M.: TRank: ranking Twitter users according to specific topics. In: Proceedings of the 12th International IEEE Consumer Communications and Networking Conference (CCNC 2015), January 2015
12. Furini, M., Montangero, M.: TSentiment: on gamifying Twitter sentiment analysis. In: Proceedings of IEEE Symposium on Computers and Communication, pp. 91–96, June 2016
13. Salomoni, P., Prandi, C., Roccetti, M., Nisi, V., Jardim Nunes, N.: Crowdsourcing urban accessibility:: some preliminary experiences with results. In: Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter, CHItaly 2015, pp. 130–133. ACM, New York (2015)
14. Mirri, S., Prandi, C., Salomoni, P.: Personalizing pedestrian accessible way-finding with mPASS. In: 2016 13th IEEE Annual Consumer Communications Networking Conference (CCNC), pp. 1119–1124, January 2016
15. Nakov, P., Rosenthal, S., Kozareva, Z., Stoyanov, V., Ritter, A., Wilson, T.: Semeval-2013 task 2: sentiment analysis in Twitter. In: Proceedings of the Seventh International Workshop on Semantic Evaluation, pp. 312–320, June 2013
16. Mitchell, L., Frank, M.R., Harris, K.D., Dodds, P.S., Danforth, C.M.: The geography of happiness: connecting Twitter sentiment and expression, demographics, and objective characteristics of place. PLoS ONE **8**(5), 1–15 (2013)
17. Lin, Y.-R.: Assessing sentiment segregation in urban communities. In: Proceedings of the International Conference on Social Computing, pp. 9:1–9:8 (2014)
18. Guo, W., Gupta, N., Pogrebna, G., Jarvis, S.: Understanding happiness in cities using Twitter: jobs, children, and transport. In: 2016 IEEE International Smart Cities Conference (ISC2), pp. 1–7, September 2016
19. Egidi, L., Furini, M.: Bringing multimedia contents into MP3 files. IEEE Commun. Mag. **43**(5), 90–97 (2005)
20. Egidi, L., Furini, M.: From digital audiobook to secure digital multimedia-book. ACM Comput. Entertain. **4**(3), 5 (2006)
21. Oliveira, E., Martins, P., Chambel, T.: Accessing movies based on emotional impact. Multimed. Syst. **19**(6), 559–576 (2013)
22. De Michele, R., Furini, M.: Understanding the city to make it smart. In: Mandler, B., et al. (eds.) IoT360 2015 Part I. LNICST, vol. 169, pp. 239–244. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47063-4_22
23. Furini, M.: Users behavior in location-aware services: digital natives vs digital immigrants. Adv. Hum.-Comput. Interact. **2014** (2014)
24. Furini, M., Tamanini, V.: Location privacy and public metadata in social media platforms: attitudes, behaviors and opinions. Multimed. Tools Appl. **74**(21), 9795–9825 (2015)
25. Xu, Z.: Trip similarity computation for context-aware travel recommendation exploiting geotagged photos. In: 2014 IEEE 30th International Conference on Data Engineering Workshops, pp. 330–334, March 2014
26. Bojic, I., Sobolevsky, S., Nizetic-Kosovic, I., Podobnik, V., Belyi, A., Ratti, C.: Sublinear scaling of country attractiveness observed from Flickr dataset. In: 2015 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), pp. 305–308, December 2015

27. Quercia, D., Schifanella, R., Aiello, L.M.: The shortest path to happiness: recommending beautiful, quiet, and happy routes in the city. In: Proceedings of the 25th ACM Conference on Hypertext and Social Media, HT 2014, pp. 116–125. ACM, New York (2014)
28. You, S., DesArmo, J., Joo, S.: Measuring happiness of US cities by mining user-generated text in Flickr.com: a pilot analysis. Proc. Am. Soc. Inf. Sci. Technol. **50**(1), 1–4 (2013)
29. Bujari, A., Ciman, M., Gaggi, O., Palazzi, C.E.: Using gamification to discover cultural heritage locations from geo-tagged photos. Pers. Ubiquit. Comput. **21**(2), 235–252 (2017)
30. Gaggi, O.: Discovering local attractions from geo-tagged photos. In: Proceedings of the 28th Annual ACM Symposium on Applied Computing, SAC 2013, pp. 730–735. ACM, New York (2013)
31. Abdullah, S., Murnane, E.L., Costa, J.M.R., Choudhury, T.: Collective smile: measuring societal happiness from geolocated images. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 361–374. ACM, New York (2015)
32. Liu, K., Motta, G., You, L., Ma, T.: A threefold similarity analysis of crowdsourcing feeds. In: 2015 International Conference on Service Science (ICSS), pp. 93–98, May 2015
33. de Oliveira, T.H.M., Painho, M.: Emotion amp; stress mapping: assembling an ambient geographic information-based methodology in order to understand smart cities. In: 2015 10th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–4, June 2015

The Use of Hashtags in TV-Shows: Analysis and Guidelines

Roberta De Michele[1] and Marco Furini[2](✉)

[1] Fondazione Marco Biagi, Universta di Modena and Reggio Emilia, 42121 Modena, Italy
roberta.demichele@unimore.it
[2] Dipartimento di Comunicazione ed Economia, Universta di Modena and Reggio Emilia, 42121 Reggio Emilia, Italy
marco.furini@unimore.it

Abstract. People watch TV with a second screen nearby and more than 30% of the time is used to perform social activities. This is why broadcasters are trying to exploit hashtags: they want people to talk about TV-shows in the social scenario. In this paper, we investigate how broadcasters use hashtags to promote their TV-shows. Indeed, we focus on 15 different TV-shows and we perform two different types of analyses: one investigates if broadcasters propose official hashtags and analyzes how these hashtags are promoted, the other investigates the characteristics of the Twitter conversations held around the considered TV-shows. Results show that broadcasters do not have a clear strategy in the social scenario: most of them do not exploit the synergy that can be created by linking the TV scenario to social and Web platforms. However, from the analysis of the Twitter conversations, we identified a successful case, whose analysis allowed us to outline some clear guidelines that broadcasters should employ when using hashtags to promote their TV-shows.

Keywords: Hashtags · TV-shows · Marketing · Twitter conversation

1 Introduction

The social scenario and the advances in Internet technologies changed the way we watch TV: in the past, the TV-set was the king of the living room and people watched TV-shows alone or with their family, but today everything has changed [1]. Young generations are no longer interested in the TV-set and migrated to streaming services that allow them to watch TV-shows on their tablets, smartphones or laptops. This migration is highlighted by a recent Nielsen Research [2]: from 2011 to 2016, 40% of young Americans moved to streaming services. Moreover, more and more people watch TV with a second screen nearby, that is with a tablet, a smartphone or a laptop at their side. For instance, recent researches show that 86% of people has a second screen in the nearby when

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 321–330, 2018.
https://doi.org/10.1007/978-3-319-76111-4_32

watching TV [3]. In particular, 81% has a smartphone in the nearby, 66% use a laptop/PC device when watching TV [4]. Furthermore, the proximity of these second devices is tempting: many studies highlighted that more than 30% of the attention time is focused on these devices [5].

In the attempt to re-catch the viewers' attention, broadcasters are trying to embrace the second screen scenario by suggesting official hashtags to create conversations around TV programs in the social scenario [6,7]. A hashtag is an alphanumerical string preceded by the hashtag special character #. Originally introduced in Twitter, the micro-blogging platform that permits to post messages up to 140 characters in length, nowadays people use hashtags in any social platform to emphasize contents and to facilitate the message classification. For instance, a user who wants to emphasize that his/her post is about "social media" may use the hashtag `#socialmedia` within the message or within the description of the posted picture; similarly, a user who wants to browse for messages that talk about social media may use the hashtag `#socialmedia` as a research keyword. In summary, hashtags provide an effective way to specify the main topic of the message and to follow conversations on specific topics (e.g., a public event, a brand, a politician, etc.).

People use hashtags to browse for any type of content (e.g., from personal feelings to public critics, from meaningless messages to important updates) posted by any type of users (e.g., from ordinary people that share personal information with their friends to important executives that talk about economics, politics and corporations) [8,9], and different studies highlighted that many conversations that are being held on Twitter concern television programs [10,11].

In this paper, we investigate how broadcasters use hashtags to promote social conversations around their TV-shows. To the best of our knowledge, different studies analyzed social media conversations to understand peoples' opinions towards TV-shows, but this is the first paper that analyzes the broadcasters' use of official hashtags. In particular, we focus our study on the Twitter platform, as this platform is one of the most used in the social scenario to talk about TV-shows [11].

Our study focuses on 15 different popular TV-shows broadcast (either through free-to-air channels, digital subscription television companies or streaming platforms) in Italy during Spring 2017. The first investigation analyzes the official Web Page and the official Twitter account of the TV-shows to see if an official hashtag is available and advertised somehow. Then, we analyze the Twitter conversations around the hashtags.

Results show that broadcasters do not have a clear strategy in the social scenario: they create hashtags, but they exploit them badly or they do not exploit them at all. From the analysis of the Twitter conversations, we identify a successful case, whose analysis allowed us to outline some clear guidelines that broadcasters should employ when using hashtags to promote the interests around their TV-shows.

This paper is organized as follows. In Sect. 2 we present a brief overview of studies that consider the relation between the social and the TV scenario. In

Sect. 3 we show details of our investigation. A set of guidelines is presented in Sect. 4, whereas Conclusions are drawn in Sect. 5.

2 Related Work

The advances in mobile and Internet technologies and the proliferation of social platforms are changing the way we live, work and have fun [8,9,12–17]. With respect to the act of watching TV programs, the use of a second device to perform social activities is becoming very popular [10].

Recent studies highlighted that Twitter is often used by people when watching TV. Indeed, Mukherjee et al. [11] analyzed the Twitter conversations for a period of time and results showed that the subject of the conversations is usually related to TV-programs and also revealed that people post messages that can be grouped into five different categories: questions, responses, referrals, broadcasts and retweet. A subsequent interesting study [18] analyzed three different social platforms (i.e., Twitter, Instagram and Tumblr) during the Superbowl. Results showed that people talk about everything: the conversations were not only related to the game, but also towards music and commercials. Holmes et al. [5] showed that when people watch TV, 30% of their attention time is focused on the smartphone.

These studies show that people pay less attention to television programs. That is, even when people sit in front of a TV screen, they are not 100% focused on the TV program: the several mobile devices they have nearby distracted them. Therefore, it does not surprise that several studies have been aimed at capturing the attention of the viewers. For instance, Meloni et al. [19] focused on the Brazilian digital television system and proposed an application for second screen devices that allows users to send messages, and immediately display them on a digital television screen. Pagno et al. [20] elaborated a set of guidelines to help developers in the use of second screens. Lee et al. [21] proposed a second screen Web service framework based on multimedia package distribution. Knowles et al. [22] proposed a second screen application for interactive TV and investigated the user experience of interactive TV with the help of second screen applications. Results showed that the perception of second screen varies and it's content specific. The investigation also revealed that quizzes are positively perceived.

In such scenario, broadcasters and advertisers are trying to exploit the social platforms to promote their products. In particular, in addition to posting activities, the use of hashtags to link the TV product (either a TV program or a TV commercial) to social activities is receiving more and more attention. Since this approach is in its early stage, we think it is necessary to analyze the current scenario in order to understand if hashtags are efficiently used, if there are successful cases and what are the main features of these successful cases.

3 Experimental Investigation

To understand how broadcasters use the social scenario, we focus our study on 15 different popular TV-shows broadcast in Italy during Spring 2017. Note that, the considered TV-shows are available through free-to-air channels, digital subscription television companies and streaming platforms. The goal of the experimental investigation is to analyze how broadcasters use hashtags to promote TV programs and to engage viewers. In the following, we first analyze the official Web Page and the official Twitter account to see if an official hashtag is available and advertised somehow. Then, we collect data from Twitter to analyze the conversations around the official hashtags that are available.

3.1 The Use of Official Hashtags

To understand the role that hashtags have in the television scenario, we analyzed the official Web page of every considered TV-show in order to see if an official hashtag is displayed on the home page; then, we analyzed whether an official account is present on the Twitter platform and if the short description of the Twitter account reports the official hashtags.

Table 1 summarizes how broadcasters employ and use hashtags to promote TV-Shows. The results highlight the broadcasters' interest towards the social scenario: every considered TV-Shows has an official hashtag. This clearly indicates that broadcasters want to stimulate social conversations.

Table 1. The 15 monitored TV-shows.

TV-show	Official hashtag	Visibility on Web page	Visibility on Twitter
13 Reasons Why	#13reasonswhy	☐	☑
American Horror Story	#americanhorrorstory	☐	☐
Breaking Bad	#breakingbad	☐	☑
Fargo	#fargo	☐	☑
Game Of Thrones	#gameofthrones	☐	☑
Grey's Anatomy	#greysanatomy	☐	☐
Modern Family	#modernfamily	☐	☐
NCIS	#ncis	☐	☑
Prison Break	#prisonbreak	☑	☑
Riverdale	#riverdale	☐	☑
Sherlock	#sherlock	☐	☑
Sons of Anarchy	#sonsofanarchy	☐	☐
The 100	#the100	☐	☐
True Blood	#trueblood	☐	☑
Twin Peaks 2017	#twinpeaks2017	☐	☑

However, the analysis also reveals the misuse in promoting the hashtags through Web pages: only one TV-show (i.e., "Prison Break") has the official hashtag displayed on its home page. The analysis also highlights that some broadcasters (5 out of 15) do not advertise the official hashtags on the official Twitter account. Figure 1 shows the descriptions of two Twitter accounts: one promotes the official hashtag and one doesn't.

Fig. 1. Description of the Twitter account. The example on the left clearly displays the hashtag, whereas the example on the right does not mention the official hashtag.

These results show that broadcasters consider both the Web and the Social scenario to promote and to engage viewers towards conversation around TV-shows. However, we notice a misuse of the social platform, as Web and social are considered two separate worlds that do not interact with each other: broadcasters did not exploit the synergy between Web and Social. Needless to say, this is a bad practice to promote TV-shows and to create engagement with the audience.

3.2 Conversations Around the Official Hashtags

To investigate whether people use the official hashtags to talk around TV-shows, we listened to the Twitter conversations centered around the 15 TV-shows (e.g., "13 Reasons Why", "American Horror Story", "Breaking Bad", etc.). After a period of 30 days, we collected around 250,000 Twitter messages.

Conversations by Hashtags. The first goal of our analysis is to understand whether the broadcaster medium (e.g., streaming, satellite, free-to-air) affects the number of conversation around TV-shows. Figure 2 displays the top ten hashtags used by people who talk about the considered TV-shows in the dataset. "Prison Break" (#prisonbreak), a five-seasons serial drama that had four seasons broadcast from 2005 to 2009, and a fifth season aired in April–May 2017, leads the group with more than 20% of the Twitter conversations; three different TV-shows (i.e., "Games of Thrones", "13 Reasons Why" and "The 100") reach percentages between 5 and 10%. It is interesting to notice that the broadcaster medium does not affect the popularity of the conversations. Indeed, "13 reasons why" is available through a streaming platform, "Game of Thrones" is broadcast through a digital subscription television company, whereas free-to-air channels broadcast "Sherlock".

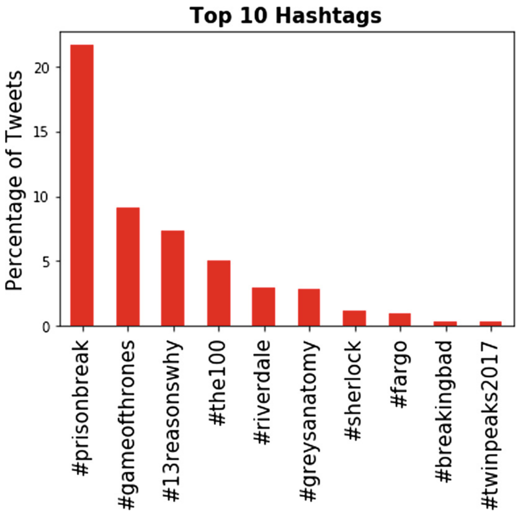


Fig. 2. Top ten hashtags used in the collected dataset.

Language. A second investigation analyzes the pervasiveness of official hashtags. Indeed, we want to understand what are the languages of the conversations. Figure 3 shows the top ten languages used in the collected dataset. It does not surprise that English is the king of the conversations (more than 65% of the conversation collected in the dataset are written in English), but it is interesting to note that people talk about TV-shows also in other languages. To deepen the analysis, Fig. 4 shows the top ten hashtags used by people who talk about TV-shows in Italian. The highest number of conversations is connected to the `#13reasonswhy` hashtag although the show title is translated from "13 Reasons Why" to "Tredici". This example highlights the pervasiveness of the official hashtag.

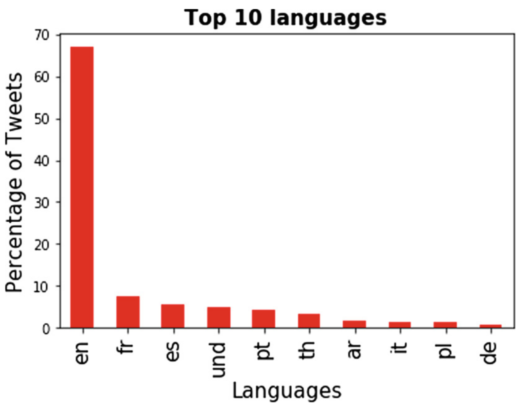


Fig. 3. Top ten languages used by people who talk about TV-shows in Twitter.

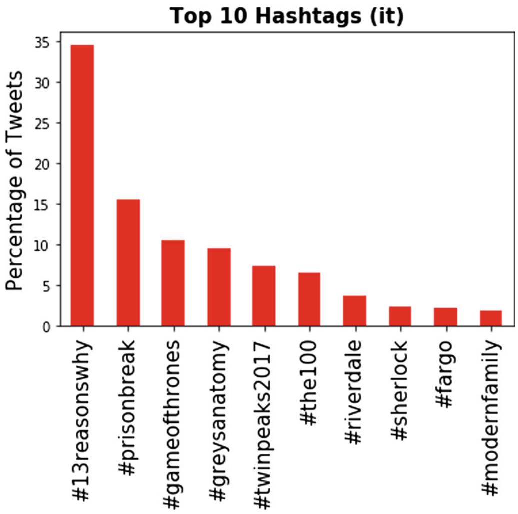


Fig. 4. Top ten hashtags (it) used by people who talk about TV-shows in Twitter.

Authors. To better understand who talk about TV-shows, Fig. 5 (left) shows the top ten authors (Twitter account) that talk about the considered TV-shows in the dataset. It is to note the low percentage reached, in general, by the accounts. There is no a specific account that leads the conversations, but these are produced by a considerable number of accounts. Among these accounts, we investigate the most retweeted ones. Indeed, in Twitter people can write their own message, but they can also contribute to the conversation by retweeting (i.e., posting again) a message wrote by someone else. The retweeting operation extend the visibility of the message. Figure 5 (right) shows the top ten most retweeted users. Some of the considered official accounts are among the most retweeted ones. `@prisonbreak` leads the group: almost 7% of the retweeted messages are written by the official account of Prison Break. Then, there are account related to TV-shows (e.g., `@13reasonszone` is the official "13 Reasons Why" fan account, `@PBWritersRoom` is the official Twitter account for the writers of the new season of Prison Break, `@PrisonBreakFive` is the French account of Prison Break) and private accounts.

Retweet Time. The `@prisonbreak` account is the most retweed one. Therefore, we focus on this account to understand the reason why people retweet its messages. First, we investigate whether people retweet messages at specific daytime. Figure 6 shows that people retweet messages at every hour of the day (with the exception of the night hours), with two peaks around 18:00–19:00 and 22:00–23:00. Likely, these are the airing time of the TV-shows. Therefore, people use to tweet while watching the TV-show. To deepen the investigation, we analyze whether the messages produced by the `@prisonbreak` account have some specific characteristics. We observed that almost every written message (i.e. 99.8%)

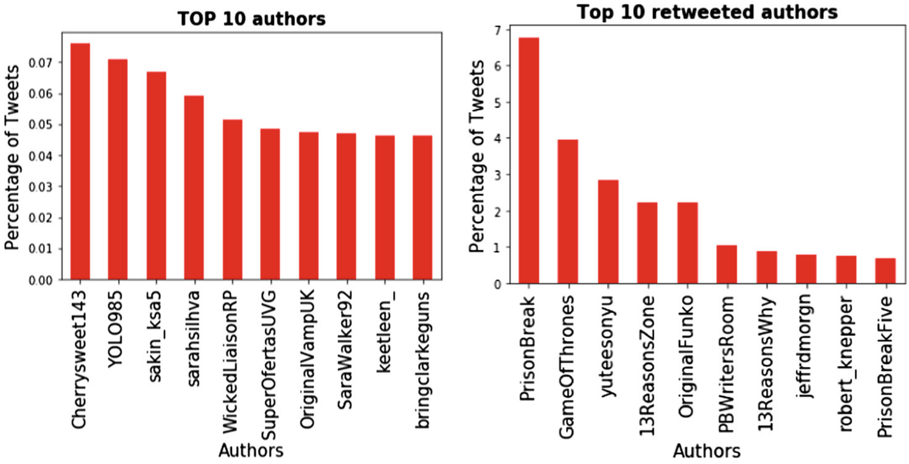


Fig. 5. Top ten authors who talk about TV-shows in the collected dataset (left) and top ten retweeted authors (right).

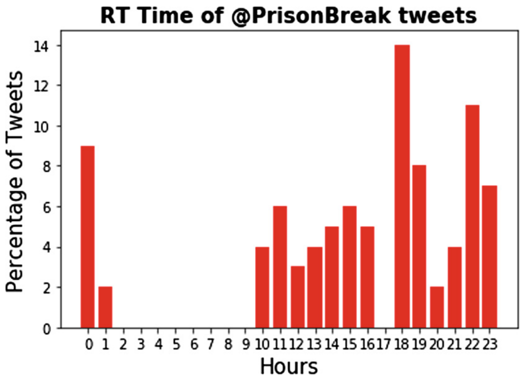


Fig. 6. Retweet time of tweets produced by the most retweeted Twitter account.

contains the official hashtag, most of them raise a question (e.g., "Missing Lincoln and Michael already?", "Love the `#PrisonBreak` journey?", etc.), almost all of them include a picture (e.g., simple image, an animated GIF or a screenshot that links to video), but what surprised us is that 58.2% of the messages asked people to retweet the message (e.g., "RETWEET if your heart broke in this moment", "RETWEET if you love the `#PrisonBreak` family!", etc.).

4 Guidelines

The analysis of the conversations observed in the dataset, allowed us to define few guidelines to help broadcaster towards the promotion of TV-shows and engagement of viewers.

- **Official Hashtag**. Create an official hashtag and an official Twitter account. Display the hashtag on the TV-show Web Page and on the description of the Twitter account. Promote the hashtag when promoting the TV-show (e.g., on flyers, newspapers/magazines ads, TV-promos, etc.).
- **Tweet**. Always use the official hashtag when writing a message on social media platforms.
- **Friendly**. Write a message as if you were talking to a person. Pose questions and express feelings.
- **Multimedia**. Enrich the message with multimedia contents. While images are ok, remember that people love animated GIF and video.
- **Ask**. Sometimes it is worth asking people to retweet a message. The `@prisonbreak` analysis shows that this technique is effective when associated with the sharing of an experience or a feeling. For example, "retweet if you feel happy", "retweet if you're watching ...", "retweet if you're playing football", etc.

5 Conclusions

In this paper, we investigated how broadcasters use hashtags to promote their TV-shows or to engage viewers. We analyzed the Web and social scenario. Results showed that broadcasters do not exploit the potential of the social scenario. Although it is clear that they understood the potential of this scenario (all the analyzed TV-shows had an official hashtag), it is also clear that ideas are very confusing. The official hashtag is almost never advertised on the Web page and it is seldom used within the description of the Twitter account. The analysis of the Twitter conversations highlighted an interesting case, whose analysis allowed us to outline some possible guidelines that broadcasters may use when using hashtags to promote their TV-shows.

References

1. De Michele, R., Furini, M.: TV commercials: improving viewers engagement through gamification and second screen. In: 2017 IEEE Symposium on Computers and Communication (ISCC), July 2017
2. Nielsen: Total audience report (2016)
3. Stenger, M.: TV goes social: the rise of the second screen (2012). http://socialnewsdaily.com/4193/tv-social-networks-use-infographic
4. Google: The New Multiscreen World: Understanding Cross-Platform Consumer Behavior (2012). https://ssl.gstatic.com/think/docs/the-new-multi-screen-world-study_research-studies.pdf
5. Holmes, M.E., Josephson, S., Carney, R.E.: Visual attention to television programs with a second-screen application. In: Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 397–400 (2012)
6. Montangero, M., Furini, M.: TRank: ranking Twitter users according to specific topics. In: 2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC), pp. 767–772, January 2015

7. Furini, M., Montangero, M.: TSentiment: on gamifying Twitter sentiment analysis. In: Proceedings of IEEE Symposium on Computers and Communication, pp. 91–96, June 2016
8. Furini, M., Tamanini, V.: Location privacy and public metadata in social media platforms: attitudes, behaviors and opinions. Multimed. Tools Appl. **74**(21), 9795–9825 (2015)
9. Furini, M.: Users behavior in location-aware services: digital natives vs digital immigrants. Adv. Hum.-Comput. Interact. **2014** (2014)
10. Courtois, C., D'heer, E.: Second screen applications and tablet users: constellation, awareness, experience, and interest. In: Proceedings of the 10th European Conference on Interactive TV and Video, EuroITV 2012, pp. 153–156. ACM, New York (2012)
11. Mukherjee, P., Wong, J.S., Jansen, B.: Patterns of social media conversations using second screens. In: Proceedings of the International BigData/SocialCom/CyberSecurity Conference (2014)
12. Bujari, A., Furini, M., Mandreoli, F., Martoglia, R., Montangero, M., Ronzani, D.: Standards, security and business models: key challenges for the IoT scenario. Mob. Netw. Appl. 1–8 (2017)
13. Egidi, L., Furini, M.: Bringing multimedia contents into MP3 files. IEEE Commun. Mag. **43**(5), 90–97 (2005)
14. Furini, M.: Mobile music distribution: a multichannel approach. Adv. Multimed. **2011** (2011)
15. Furini, M.: An architecture to easily produce adventure and movie games for the mobile scenario. ACM Comput. Entertain. **6**(2), 1–16 (2008)
16. Bujari, A., Furini, M., Laina, N.: On using cashtags to predict companies stock trends. In: 2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC), pp. 25–28, January 2017
17. Furini, M., Mandreoli, F., Martoglia, R., Montangero, M.: IoT: science fiction or real revolution? In: Gaggi, O., Manzoni, P., Palazzi, C., Bujari, A., Marquez-Barja, J.M. (eds.) GOODTECHS 2016. LNICST, vol. 195, pp. 96–105. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61949-1_11
18. Mukherjee, P., Jansen, B.J.: Second screen interaction analysis for IRL events: phase-category investigation of the super bowl 2015 social soundtrack. In: 2016 7th International Conference on Information and Communication Systems (ICICS), pp. 13–19, April 2016
19. Meloni, L.F.P., Costa, G.C., Kobayashi, G., Kurashima, C.S.: Implementation of chat application for Ginga middleware technology using second screen. In: 2016 IEEE International Symposium on Consumer Electronics (ISCE), pp. 43–44, September 2016
20. Pagno, B., Costa, D., Guedes, L., Freitas, C.D.S., Nedel, L.: Guidelines for designing dynamic applications with second screen. In: 2015 XVII Symposium on Virtual and Augmented Reality, pp. 42–51, May 2015
21. Lee, S., Yoo, J., Han, G.: A second screen web service framework based on multimedia package distribution. In: 2015 IEEE International Conference on Consumer Electronics (ICCE), pp. 42–45, January 2015
22. Mu, M., Knowles, W., Sani, Y., Mauthe, A., Race, N.: Improving interactive TV experience using second screen mobile applications. In: 2015 IEEE International Symposium on Multimedia (ISM), pp. 373–376, December 2015

A Comparison of a Smartphone App and a Wrist-Worn Fitness Tracker for Self-monitoring of Physical Activity by Older and Younger Users

Rebekka Kupffer, Melanie Wutzler, Josef F. Krems, and Georg Jahn(✉)

Institute of Psychology, Chemnitz University of Technology, Wilhelm-Raabe-Str. 43, 09120 Chemnitz, Germany
{rebekka.kupffer,melanie.wutzler}@s2014.tu-chemnitz.de, {josef.krems,georg.jahn}@psychologie.tu-chemnitz.de

Abstract. Wearables capable of monitoring steps are important elements of behavior change interventions to increase physical activity. For intervention studies, there is currently the choice between smartphone apps and fitness trackers for self-monitoring daily steps. We report results from a pilot study, in which younger and older participants experienced both types of devices, rated usability, and performed usability tasks. The fitness tracker, which was operated with a single touch-sensitive button, proved advantageous in subjective and objective usability. The discussion includes further aspects of the choice between smartphone apps and fitness trackers for use in interventions.

Keywords: Mobile health · Wearables · Quantifying self · Behavior change Pedometer app · Accelerometry

1 Introduction

Increasing physical activity is of foremost importance in improving public health. Self-monitoring is a main component of behavior change interventions to increase physical activity [1] and wearable devices that record physical activity support self-monitoring very effectively. They capture activity more accurately than subjective estimates, which require effortful awareness and are distorted by memory lapses, imprecision, and motivated judgment. Before smartphones and fitness trackers became widely available, pedometers have proven effective in increasing daily physical activity quantified by step counts. In studies, in which participants recorded their daily step counts in diaries and had been assigned a goal of daily steps to strive for, the increase achieved by pedometers, diaries, and goal setting in groups of sedentary adults was about 2000 steps per day [2]. The effect of such an intervention is likely attenuated if

R. Kupffer and M. Wutzler have contributed equally.

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 331–341, 2018.
https://doi.org/10.1007/978-3-319-76111-4_33

individuals are not participating in a scientific study and long-term effects are still unclear. Moreover, the achievable increase in physical activity is likely smaller for older adults with impaired mobility. Nonetheless, technological support for self-monitoring is an important element of behavior change interventions even in older age groups.

Today, fitness trackers and smartphone apps are available for self-monitoring of physical activity. They afford potentially powerful additional functionality for achieving behavior change (e.g., adaptive goal setting, individualized feedback, prompts, rewards, suggestions for activity, social networking, and social support), however, particularly for older target groups, usability will critically impact acceptance and consequently effectiveness. Several aspects of usability differ between the two types of wearables and need to be considered when deciding which type should be employed in a behavior change intervention with older adults.

First of all, individuals who are already using a smartphone are likely familiar with its basic functions, can power it on and off, can activate the display, can charge the smartphone, can access apps, and can switch between apps. All these necessary operations can be serious obstacles for less experienced individuals. For older users who do not own a smartphone and lack experience with computers and digital wearables, a standalone fitness tracker may be better suited for simple self-monitoring of physical activity if it does not pose serious usability problems. Presumably, a simple standalone fitness tracker will be better accepted and used with less problems than a smartphone app.

For self-monitoring, it is important that the devices can be used without assistance. Many fitness trackers require to be used with a smartphone or a computer for being set up and for accessing activity data as statistics over multiple days. Setting up the fitness tracker could be performed by a second person, but day to day use should not require assistance by a second person. Putting the fitness tracker on and off, charging, and accessing recorded activity data all have to be managed by the person who uses the device for self-monitoring. We test for both device types how capable older users are of these tasks that have to be performed in daily use.

For exploring, whether to employ a smartphone app or a fitness tracker in behavior change interventions with older individuals, and to study the usability of examples of both device types, we enrolled participants in a study that consisted of three phases each lasting 5 days. Accelerometry was used in all three phases for a reference measurement of physical activity. In the first phase, participants just wore the accelerometer, estimated their daily steps, and noted their estimate in a diary. For the second and third phase, they were asked to strive for a daily step goal of 7000 steps and used either a smartphone app or a wrist-worn fitness tracker for self-monitoring. They kept a diary of the daily step counts. Usability ratings and usability tests were scheduled before and after the participants had gained experience in using the devices for self-monitoring.

2 Method

2.1 Participants

The study included 20 participants in an older and a younger age group. The 13 older participants (7 female, 6 male) were between 65 and 81 years of age and were recruited in a lecture for senior citizens, on a public square in the city of Chemnitz, and from a participant database. The 7 younger participants (4 female, 3 male) were between 21 and 30 years of age and were recruited via a student mailing list or at Chemnitz University of Technology. Demographic information about the participant groups is shown in Table 1. Participants received 40 Euros as compensation. Three students chose to collect participant hours as part of a curricular requirement instead.

Table 1. Demographic information on the older (>65 years) and younger (<30 years) participant groups.

	Older (N = 13)	Younger (N = 7)
Age in years	70.03 (4.48)	24.87 (3.38)
BMI in kg/m^2	26.59 (2.71)	21.87 (1.64)
CPQ-score	18.76 (6.15)	26.99 (2.86)
SPQ-score	15.27 (5.29)	23.08 (4.49)

Note. Means and standard deviations (in brackets). BMI: body mass index; CPQ: Computer Proficiency Questionnaire; SPQ: Smartphone Proficiency Questionnaire

2.2 Devices

Smartphone and Pedometer App. When physical activity was self-monitored with a smartphone app, participants used Google Nexus 5 Android smartphones with the "Pacer" pedometer app. The app was installed and set up before the participants received the smartphones. In March 2017, Pacer was the best-rated pedometer app (4.6/5 stars) in the German Android app-store. The app is always running in the background. Therefore, it is not necessary to start a measurement before going for a walk. The step count for the current day is shown on the Android homescreen. Additional information can be accessed when the app is opened (see Fig. 1). GPS tracking has to be activated in order to pace the steps made.

Fig. 1. A screenshot from the pedometer smartphone app (Pacer), the fitness tracker (Mi Band 2), and the toggling cycle for display content on the fitness tracker.

The smartphone was carried in a waistbelt that also contained the accelerometer (see Fig. 2). Participants were instructed to wear the smartphone on the right hip.

Fig. 2. Waistbelt and movisens Move II accelerometer.

Wrist-Worn Fitness Tracker. When physical activity was self-monitored with a fitness tracker, participants used the Mi Band 2 (Xiaomi). The fitness tracker is usually worn in a wristband and shows information on an OLED display that is activated and toggled by a single touch-sensitive button (see Fig. 2). By default, three screens can be cycled through: The current time is shown when the display is activated. The second touch switches to the step count, which is automatically reset at midnight. A third touch switches to measuring the heart rate. A pounding heart symbol is shown while the heart rate is measured with an optical sensor, then the heart rate is shown. The display turns off automatically and starts with showing the time when activated again.

The Mi Band 2 has to be set up with a smartphone app that is available for iOS and Android. The fitness tracker was set up and configured before the participants received it. They did not use the fitness tracker together with a smartphone and thus, the additional features available in the app such as a history of daily step counts, energy consumption, covered distance, sleep quality etc. were not accessed. Participants were instructed to wear the fitness tracker on the left wrist.

Accelerometer. For baseline and reference measurements of step counts, each participant wore a movisens Move II activity tracker (movisens, Karlsruhe, Germany) (see Fig. 2). It combines a three-axis accelerometer, a barometric altimeter, and a temperature sensor. The raw data can be processed by the DataAnalyzer software from movisens to generate activity reports including step counts. The participants were instructed to wear the accelerometer in the waistbelt on the right hip.

2.3 Materials and Tasks

A questionnaire on computer proficiency, a questionnaire on smartphone proficiency, and a brief usability scale to collect subjective ratings were employed. The Computer Proficiency Questionnaire (CPQ) [3] asks for subjective ratings on a 5-point scale of the ability to perform tasks with a computer (basic tasks, printing, communication, calendar functions, using the internet, entertainment). The CPQ-score varies between 6 and 30 and higher values indicate higher proficiency. We used a German translation of the CPQ and prepared a smartphone proficiency questionnaire (SPQ) consisting of 13 items by rephrasing appropriate items of the CPQ. The SPQ-score varies between 5 and 25.

For subjective ratings of usability, the System Usability Scale (SUS) [4] was used. It collects ratings on a 5-point scale for 10 items, which are then converted to a score between 0 and 100. Higher scores indicate better usability.

A different set of tasks was put together for each device to test usability (see Table 2). Participants' hands and the devices were recorded on video while the tasks were performed. Task times (efficiency) and error frequencies (effectiveness) were coded from the videos.

Table 2. Usability tasks performed with the smartphone app and with the fitness tracker.

Smartphone app	Fitness tracker
1. Switch the display on	1. Put the fitness tracker in the wristband
2. Search for the pedometer app and open the app	2. Put the wristband on
3. Walk a round in the room and check the step count (repeated two times)	3. Switch the display on
4. Search for a screen that shows yesterday's step count	4. Search for the step count
5. Search for a screen that shows the total distance recorded in the app	5. Walk a round in the room and check the step count (repeated two times)
	6. Check your heart rate
	7. Take off the wristband

2.4 Procedure

The study encompassed four appointments at the university and three 5-day intervals of activity tracking as shown in Fig. 3. Before the first appointment, participants were informed about the purpose and procedure of the study and provided informed consent.

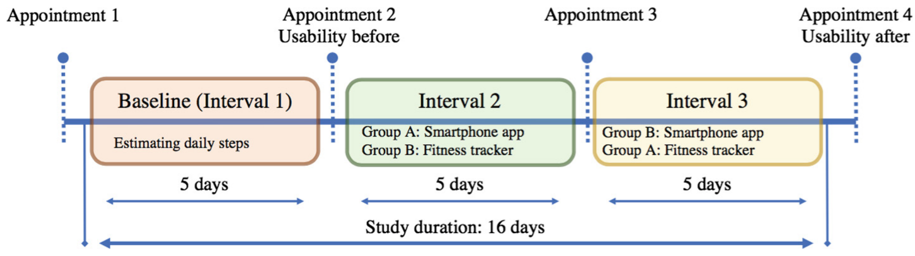


Fig. 3. Procedure with scheduling of appointments and intervals of activity tracking. Usability tests and usability ratings for both device types took place at appointments 2 (before using the devices for self-monitoring) and 4 (after self-monitoring).

At the first appointment, the tests of physical activity were conducted, participants filled in a questionnaire about demographic information, the CPQ, and the SPQ. They were given a diary and were instructed to fill in at the end of each day the number of steps they had taken, information about the weather (good, neutral, bad), and their mood (good, neutral, bad). The participants also received a short manual on how to use and charge the devices.

For the first interval of activity tracking (baseline), the participants received the movisens accelerometer and the waistbelt. In the first five days of activity tracking, they were supposed to estimate how many steps they had taken and to note the estimated step count in the diary every night before going to bed.

At the second appointment, participants performed the usability tasks first with the fitness tracker and then with the smartphone. They were instructed to try to find the solutions by their own, without any help. When they indicated that they needed help or when they could not complete a task, the experimenter demonstrated the task and the participant tried again. Then, they filled in the SUS whereby each question was asked for the fitness tracker at first, and then for the smartphone app. This was followed by some open questions.

For the second interval of activity tracking, participants received an activity goal of a minimum of 7000 steps per day and their first activity tracker. Half of the participants (group A) started with the smartphone app and the remaining half (group B) with the fitness tracker. Participants were instructed to document the step counts indicated by the respective device in their diaries each day before going to bed together with information on the weather and their mood for this day.

At the third appointment, participants in group B returned the fitness trackers, which were checked, charged, reset, and then given to participants in group A. Participants in group B used the smartphone app in the subsequent activity tracking interval. The third interval of activity tracking proceeded as the second interval. The instructions (activity goal of 7000 steps per day minimum and diary notes of step count, weather, and mood) remained the same. At the fourth and final appointment participants returned the devices and the usability-tests from the second appointment were repeated with the same instructions, followed by the SUS ratings and an open interview.

3 Results

3.1 Step Counts

All 20 participants completed the study resulting in 300 days of activity tracking, of which accelerometer recordings of eight days (2.7%) are missing because of technical failure. Participants' mean estimates of daily step counts can be compared to the mean daily step count recorded by the accelerometer for the five days in the first interval of activity tracking (baseline). Both means are shown separately for each participant in Fig. 4. For about half of the participants the mean daily estimate was clearly off the actual mean daily step count and more participants underestimated than overestimated their mean step count.

Figure 5 shows boxplots of mean daily step counts by type of tracker and age group, and by weather and gender. Mean step counts were higher for males than for females, higher at days with positive than negative weather ratings, and higher at days with positive than negative mood ratings. The type of fitness tracker and the age of participants

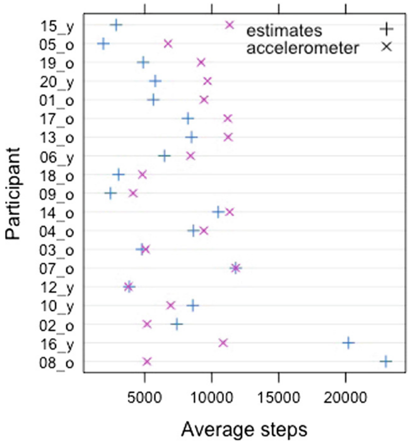


Fig. 4. Means of estimated and accelerometer daily step counts in the first interval of activity tracking (baseline) separately for each of 19 participants (missing data for 11_y) ordered by the difference between actual and estimated mean step counts.

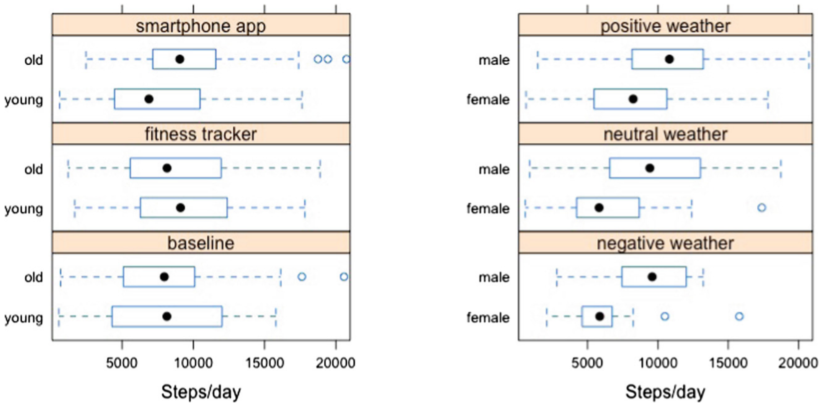


Fig. 5. Boxplots of mean accelerometer step counts by type of tracker and age group (left), and by weather rating and sex (right).

had no consistent effects on mean step counts. Linear mixed effects modeling of step count data from all three intervals of activity tracking confirmed that sex, weather, and mood but neither device type nor age were significant predictors of step counts.

3.2 Usability Tasks and Usability Ratings

All tasks and the SUS ratings were performed twice by all participants, once *before* they used the devices for self-monitoring (at appointment 2) and once *after* they had used the devices for self-monitoring. The tasks differed for the two devices, thus performance data were analyzed and are reported separately for each device.

Task effectiveness was quantified by the average error frequency for each device, which indicates how often participants could not complete tasks or asked for help. Task efficiency was quantified by the total time needed to perform the tasks.

Table 3 presents mean error frequencies and mean total times for the smartphone app and for the fitness tracker, before and after self-monitoring, and separately for the younger and the older age group. Both variables indicate better performance at the second test session (after) and for the younger age group reflecting practice and age effects. There was no difference between the age groups in the mean error frequency with the fitness tracker at the second session (both at zero) (Table 3).

Table 3. Mean error frequencies and mean total times for the usability tasks performed with the smartphone app (left) and the different usability tasks performed with the fitness tracker (right) before and after self-monitoring and separately for the younger and the older age group.

	Smartphone app		Fitness tracker	
	Young	Old	Young	Old
Error frequency				
Before	4.42 (2.86)	19.27 (3.59)	1.59 (1.58)	4.27 (2.00)
After	0.00 (0.00)	13.19 (5.23)	0.00 (0.00)	0.00 (0.00)
Time (sec)				
Before	155.86 (18.40)	339.14 (49.64)	77.14 (10.29)	129.21 (17.42)
After	65.62 (11.59)	268.33 (68.07)	40.71 (3.67)	60.82 (4.40)

Note. Means and standard errors (in brackets). Because different tasks were performed, data cannot be compared directly between devices.

Figure 6 shows mean SUS-scores with standard errors for the fitness tracker and the smartphone app before and after using the devices for self-monitoring, separately for

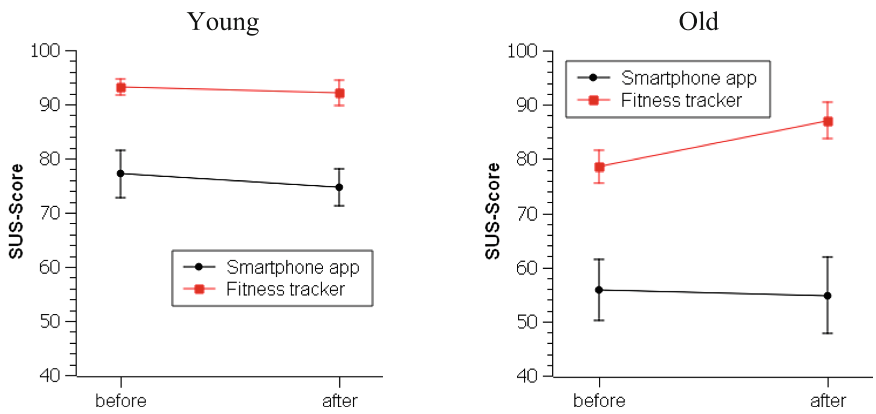


Fig. 6. Mean SUS-scores with standard errors for the fitness tracker and the smartphone app before and after using the devices for self-monitoring and separately for the younger (left) and the older (right) age group.

the younger and the older age group. Usability was rated higher for the fitness tracker at both sessions in both age groups and the difference in mean SUS-scores slightly increased from before to after in the older age group.

Furthermore, the influence of computer proficiency, smartphone proficiency, and age on task time was examined by separate linear regressions. These regressions were computed only for the older group (see Table 4) because there was hardly any variance in the CPQ and SPQ-scores in the younger group. Self-rated smartphone proficiency (SPQ) was a better predictor of task time than self-rated computer proficiency with both devices and importantly, smartphone proficiency was less influential for the time required for fitness tracker tasks. For the smartphone app tasks, smartphone proficiency was a better predictor of task time than age in the older group.

Table 4. Linear regressions predicting the total time required for performing usability tasks in the older group by self-rated computer proficiency, smartphone proficiency, and age.

	Smartphone app			Fitness tracker		
	β	*SE*	R^2	β	*SE*	R^2
CPQ	−0.06	8.75	.00	−0.25	2.97	.06
SPQ	−0.57	9.04	.33	−0.36	3.43	.13
Age	0.44	10.58	.19	0.40	3.79	.16

Note. No direct comparisons are possible between the results for the smartphone app and the fitness tracker because different usability tasks were performed.

4 Discussion

Two devices for self-monitoring daily steps were compared in a pilot study for an informed decision about which to employ in an intervention to increase the physical activity of older adults. Younger and older adults used both a smartphone app and a wrist-worn fitness tracker. The fitness tracker was rated more favorably than the smartphone app by older and younger participants. This usability advantage of a simple wrist-worn fitness-tracker over a smartphone pedometer app was also confirmed by a smaller age effect on usability task performance particularly after some practice.

The participants used the devices while striving for a goal of 7000 steps per day after they had already estimated daily step counts during a baseline interval. Comparisons and analyses of step counts in the present study should be interpreted cautiously because the sample size was small and activity was recorded for a few days only. In the present study, there was no increase in physical activity from the baseline but results from larger previous studies suggest that both devices could increase physical activity in behavior change interventions [1]. Baseline step counts could have been increased already by self-monitoring in the form of estimating daily steps and noting the estimates in a diary. The comparison of estimated with accelerometer step counts clearly and unsurprisingly shows that any technical support improves self-monitoring considerably.

The clear usability advantage of the fitness tracker once again demonstrates that reduced complexity (one button, no choice reaction) increases accessibility and acceptance. Notably, this result is in contrast with the conclusion in a previous study [5] comparing a smartphone app and a fitness tracker, in which participants had experienced only their preferred device and the smartphone app was rated more favorably. In the present study, 19 of the 20 participants stated in the final interview that they would prefer the fitness tracker for everyday use. This needs to be qualified a bit because some also noted that it was inconvenient to carry the smartphone for the sole purpose of tracking steps or as a second smartphone. Of course, this disadvantage would dissolve if an app would be used on a smartphone that is carried anyway.

Several differences between the device types that were not the focus of the present study deserve to be mentioned because they may be important if devices are chosen for interventions. If a person already uses a smartphone, self-monitoring with an app on one's own device is possible and experience with the smartphone and its operating system may help in using the app. Ideally, the device for self-monitoring would be worn continuously while activity occurs. A smartphone is likely carried less continuously than a wrist-worn fitness tracker. To really carry it continuously, the smartphone has to be charged over night or during other intervals without activity. Smartphones have to be charged more often than simple fitness trackers. On the other hand, putting on the fitness tracker may be forgotten. With a smartphone that is used for various purposes it is less likely to miss extended intervals of activity. How a smartphone is carried varies more than how a wrist-worn fitness tracker is carried. Certain ways of carrying the smartphone may result in more accurate step counts than others, however, studies examining the accuracy of step counts for recent smartphone apps arrived at quite positive accuracy evaluations. In the present study, the smartphone was carried in a belt. A wrist-worn fitness tracker more easily than a smartphone can be equipped with sensors that can pick up physiological signals. Measuring the heart rate may be a valuable functionality in interventions targeting physical activity of older individuals because the interventions may aim at activity of a certain intensity and the heart rate may be used as feedback for ensuring activity at the intended intensity level (e.g., walking at the right speed to be physically active with individually moderate intensity). Overall, a simple fitness tracker seems to be the better choice in intervention studies that include smartphone novices.

References

1. Sullivan, A.N., Lachman, M.E.: Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. Front. Public Health **4**, 289 (2017)
2. Bravata, D.M., Smith-Spangler, C., Sundaram, V., Gienger, A.L., Lin, N., Lewis, R., Stave, C.D., Olkin, I., Sirard, J.R.: Using pedometers to increase physical activity and improve health: a systematic review. JAMA **298**(19), 2296–2304 (2007)

3. Boot, W.R., Charness, N., Czaja, S.J., Sharit, J., Rogers, W.A., Fisk, A.D., Mitzner, T., Lee, C.C., Nair, S.: Computer proficiency questionnaire: assessing low and high computer proficient seniors. Gerontologist **55**(3), 404–411 (2015)
4. Brooke, J.: SUS: a quick and dirty usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (eds.) Usability Evaluation in Industry, pp. 189–194. Taylor & Francis, London (1996)
5. Fong, S.S.M., Ng, S.S.M., Cheng, Y.T.Y., Zhang, J., Chung, L.M.Y., Chow, G.C.C., Chat, Y.T.C., Chan, I.K.Y., MacFarlane, D.J.: Comparison between smartphone pedometer applications and traditional pedometers for improving physical activity and body mass index in community-dwelling older adults. J. Phys. Ther. Sci. **28**(5), 1651–1656 (2016)

On the Equivalence Between Community Discovery and Clustering

Riccardo Guidotti[1](✉) and Michele Coscia[2,3]

[1] KDDLab, ISTI-CNR, Pisa, Italy
riccardo.guidotti@isti.cnr.it
[2] CID, Harvard University, Cambridge, USA
michele_coscia@hks.harvard.edu
[3] Naxys, University of Namur, Namur, Belgium

Abstract. Clustering is the subset of data mining techniques used to agnostically classify entities by looking at their attributes. Clustering algorithms specialized to deal with complex networks are called *community discovery*. Notwithstanding their common objectives, there are crucial assumptions in community discovery – edge sparsity and only one node type, among others – which makes its mapping to clustering non trivial. In this paper, we propose a community discovery to clustering mapping, by focusing on transactional data clustering. We represent a network as a transactional dataset, and we find communities by grouping nodes with common items (neighbors) in their baskets (neighbor lists). By comparing our results with ground truth communities and state of the art community discovery methods, we show that transactional clustering algorithms are a feasible alternative to community discovery, and that a complete mapping of the two problems is possible.

Keywords: Community discovery · Clustering · Problems equivalence

1 Introduction

Data mining is a collection of techniques developed to extract latent knowledge from data, usually with few *a priori* assumptions. One of the most common tasks in data mining is the grouping of entities according to their similarity. We call this task *clustering*. Clustering allows us to understand when to treat different objects as part of the same class, whether we are planning a marketing campaign targeting homogeneous customers, or suggesting similar movies or books to similar audiences in a collaborative filtering entertainment platform.

Given our ever more interconnected society, complex networks emerged as one of the most successful models to mine social data. In this scenario, clustering proved to be one of the most studied problems, under the name of *community discovery*. Just like in clustering, the aim of community discovery is to group entities – in this case nodes. The main difference between clustering and

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 342–352, 2018.
https://doi.org/10.1007/978-3-319-76111-4_34

community discovery is that they are defined on data represented in a different way. However, a number of researchers pointed out that they share many commonalities [3].

Still, one cannot trivially apply community discovery to solve clustering and vice versa. Notwithstanding their similarities, there are important differences in the approaches. Community discovery assumes sparse connections [3], while clustering can work with dense datasetsc [17]; in clustering we usually deal with attributes with multiple types, while community discovery usually deals with a single attribute type – edges – often binary, in the case of unweighted networks.

In this paper, we propose a community discovery → clustering mapping. To do so, we narrow down to the specific clustering problem defined for transactional data: the task of grouping baskets sharing common subsets of items. This allows us to represent a network as a transactional dataset, in which each node is a basket containing its ego-network. Then, by applying transactional clustering algorithms, we group nodes sharing common neighbors.

We perform this operation on networks with ground truth communities, i.e., indicating the real communities of the nodes. We compare the result of transactional clustering algorithms with state of the art community discovery. We find that transactional clustering provides similar results when compared with it.

Our results have a number of consequences. They suggest that the community discovery → clustering mapping is feasible: notwithstanding their differences, it is possible to map one problem into the other. This enlarges the toolbox of researches looking for network communities: by applying our mapping, they could tap into a whole new set of algorithms that previously were not considered for their problem. Our results open the way to a complete community discovery ↔ clustering, which still eludes us, but for which we provide a roadmap.

2 Related Work

In the literature we find a limited number of works on the relationship between community discovery and clustering of transactional data.

In [9] the authors offer an in-depth comparative review of methods for clustering directed networks along two orthogonal classifications: one focuses on methodological principles, while the other classifies methods from the viewpoint of their assumptions about what a good cluster in a directed network is.

The authors of [10] try to discover communities in social networks through a spectral clustering method that makes full use of the network features. The core members are used for mining communities and it also exploits Page Rank for the spectral clustering initialization, to avoid the sensitivity to the initial centroids.

In [20] the authors face the community discovery problem by using *Non-negative Matrix Factorization* (*NMF*). NMF is a powerful interpretable tool with a close relationship with clustering methods. They target different types of networks (undirected, directed and compound) using three different NMF techniques: Symmetric NMF, Asymmetric NMF and Joint NMF.

In [14] the authors present a multi-level algorithm for graph clustering using simulating stochastic flows. The graph is coarsened to a manageable size, and then the algorithm performs on it a small number of iterations of flow simulations. The graph is successively refined with flows from the previous graph used as initializations for brief flow simulations on each of the intermediate graphs. Finally, the high-flow regions of the final refined graph are clustered together, with regions without any flow forming the natural boundaries of the clusters.

In [18] we have a clustering method for learning a binomial mixture model in RNA graphs, to process text represented as graphs in the field of bioinformatics. Textual data is analyzed also in [8], where the authors study the relations of text with topics and communities. They propose a hierarchical community model distinguishing community cores from affiliated members. The method identifies potential cores as seeds of communities through relation analysis and eliminates the influence of initial parameters through an attribute-based core merge process.

The authors of [7] present a clustering algorithm for reducing the visual complexity of a large network by temporarily replacing a set of nodes in clusters with abstract nodes. The approach defines a node similarity measure used to build a similarity matrix. The linkage pattern of the graph is thus encoded into the similarity matrix, and then one obtains the hierarchical abstraction of densely linked subgraphs by applying the k-means algorithm [17] to the matrix.

Even though in a certain sense all the works described above consider simultaneously both community discovery and clustering, this is done only for practical purposes. Here, we consider this mapping in a principled way, not to find a good heuristic, but to overcome the differences between the two problems and using existing techniques in one problem to solve the other.

3 Problem Definition

In this section we first define the problem of community discovery in complex networks. Then, we define the problem of clustering transactional data. Finally, we offer a mapping of the two, suggesting possible avenues for their equivalence.

Community Discovery ($\mathcal{CD}$). Let $G = (V, E)$ be a graph, where V is the set of nodes in the graph, and $E \subset V \times V$ is the set of edges, or node pairs. In $\mathcal{CD}$, we start from the assumption that membership to E is not random, but it follows an unknown function f. Function f governs the probability of two nodes i and j to be connected. There are different ways to model f, each one of them implying a different definition of *what a community is* [3].

In the following, we adopt the branch of community discovery assuming that nodes have unobserved attributes from a set A, and the more attributes two nodes have in common the more likely they will be part of the same community[1]. If $P = \{p_1, \ldots p_K\}$, with $p \subseteq V$, is a node partition extracted from the

[1] Note that this is far from being an unproblematic definition [11], but it will do for the scope of this paper.

space of all possible partitions of V, then the aim of community discovery is to find $argmax_P \gamma(A, P)$, where γ is a function comparing P to the actual node attributes, for instance their mutual information. Note that, if for nodes i and j $a_i = a_j$, then $f(i, j) \sim 1$. Therefore, a good partitioning P of the nodes should guarantee that $P \sim f$. In other words, if $p \in P$ and if $i, j \in p$, i.e., i and j are part of the same community p, then $(i, j) \in E$ with high likelihood.

Transactional Clustering Problem ($\mathcal{TC}$). Moving to $\mathcal{TC}$, let $B = \{b_1, \dots, b_N\}$ be a set of N baskets (transactions) and $I = \{i_1, \dots, i_D\}$ a set of D items. A basket b_i is a subset of items such that $\varnothing \subset b_i \subseteq I$. The transactional clustering problem consists in finding the partition P of B such that $argmax_P \pi(B, P)$, where π considers the average size of all baskets pairwise intersection $b_i \cap b_j$. Baskets b_i, b_j belong to the same basket partition $p \in P$ if they have a large item overlap. Thus, a good partition P puts together baskets which all share the largest possible subset of items I.

Problems Equivalence. The problem definitions of community discovery $\mathcal{CD}$ and transactional clustering $\mathcal{TC}$ share a similar structure, starting from different input types. We propose a $\mathcal{CD} \rightarrow \mathcal{TC}$ mapping between the two problems and a proper mechanism to transform the $\mathcal{CD}$ input into a $\mathcal{TC}$ input. The reasons why we do not propose $\mathcal{CD} \leftrightarrow \mathcal{TC}$ are discussed at the end of the section.

Let us consider a function Θ. Θ takes as input a graph G and a node $i \in V$, and it returns i and all $j \in V$ such that $(i, j) \in E$. This is the list of direct neighbors of i, or the set of nodes part of i's ego-network. Ego-networks have been successfully adopted to solve the $\mathcal{CD}$ problem [5]. In practice, the output of $\Theta(G, i)$ for all $i \in V$ can be seen as the input set B of baskets in the $\mathcal{TC}$ problem, containing as items the nodes ids. Figure 1 shows the transformation of a network into a set of transactions passing through an adjacency matrix. Indeed, a third way to formulate this problem is the factorization of the adjacency matrix [20].

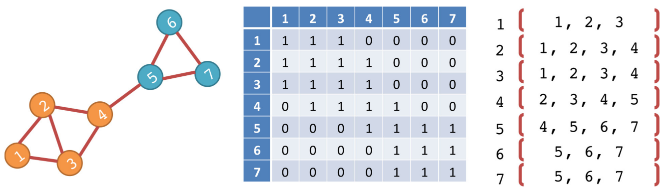


Fig. 1. Transformation of a network (left) into a set of transactions (one for each node, right), passing through the adjacency matrix (middle) of the network.

Notice that, given $\Theta(G, i) = b_i$, then $N = D = |V|$, i.e., the number of baskets is equivalent to the number of items in I and to the number of nodes in G. Moreover, we have that if $i \in V$, then $i \in I$ too, i.e., each node is an item that can be contained in a basket b_j. It follows that $I = V$.

If a partition P satisfies the objective of $\mathcal{TC}$, then all b_i and b_j in the same cluster $p \in P$ share a lot of items from I. Since $I = V$ and $i, j \in V$, then we can rewrite the previous sentence as: if P satisfies the objective of $\mathcal{TC}$, this means that i and j have a lot of common neighbors k, i.e., $f(i,k) \sim 1$ and $f(j,k) \sim 1$. In turn, this means that it is likely that both $a_i = a_k$ and $a_j = a_k$. From the transitive property of equivalence, it follows that $a_i = a_j$, which implies $f(i,j) \sim 1$. Thus, P satisfies $\mathcal{CD}$ by grouping i and j together.

To sum up: if P satisfies the objective of $\mathcal{TC}$, then it will also satisfy $\Theta(\mathcal{CD})$. As a result of this transformation, we can use algorithms developed to solve the transactional clustering problem to solve the community discovery problem.

We cannot apply Θ to B and map $\mathcal{TC}$ into $\mathcal{CD}$ because the concept of ego-network makes no sense in a transactional dataset. To perform such mapping, we need a function $\Omega = \Theta^{-1}$ that is the functional inverse of Θ. However, while Θ is a trivial operation on a graph [5], Ω is not the same on transactional data.

First, Θ has the property of automatically imposing $I = V$, which means that a basket b_i is a set of basket ids. This equivalence is broken in most transactional datasets: I is not the set of basket ids. Therefore, if Ω transforms $b_i = \{i_1, i_2, i_3\}$ into a network, it will result in a bipartite one, with two node classes: B and I. Bipartite community discovery is a different problem from traditional community discovery [3]: for instance, for two nodes $i, j \in B$ classified in the same p, $f(i,j) = 0$, because in a bipartite network B nodes connect exclusively to I nodes.

Even if we assume $I = V$, in a transactional dataset it is likely that any pair of items will appear at least once in a basket. If $I = V$, a naive Ω will result in a complete network, where $E = V \times V$. In such scenario, algorithms solving $\mathcal{CD}$ break. Ω would have to incorporate a criterion for discarding weak edges, which is far from being a solved problem, where naive strategies of filtering low weights fail [4]. For these reasons, we do not pursue the $\mathcal{CD} \leftrightarrow \mathcal{TC}$ mapping, and we focus on the $\Theta(\mathcal{CD}) \rightarrow \mathcal{TC}$ mapping for the rest of the paper.

4 Methods

Here, we describe some state-of-the-art methods addressing the community discovery and clustering of transactional data problems. We highlight the similarities and differences between the two classes of algorithms.

4.1 Community Discovery Methods

Louvain [1] is a bottom-up community discovery method maximizing the *modularity* to discover the node partition. It detects which node-community assignment would maximize the modularity, until there is no possible improvement.

LabProp [12] is a community detection method based on *label propagation*. It follows a bottom-up strategy by initially assigning a different label to each node. Then it iteratively re-labels each node with the label attached to the majority of its neighbors. This process is repeated until no further re-labeling is performed. In *LabProp*, there is not a measure to maximize, only a condition to respect.

Infomap [13] is a bottom-up community discovery approach which exploits a random walker as a proxy of information flow and then tries to minimize the *map equation*, which encodes the node identifiers initially assigned.

4.2 Transactional Clustering Methods

Clope [22] uses a global *profit* function that tries to increase the intra-cluster overlapping of transaction items by increasing the height-to-width ratio of the cluster items frequency histogram. *Clope* scans the transactions and it evaluates the profit in placing a basket in a cluster, or in creating a new one. The procedure ends when no baskets are moved from a cluster to another. *Clope* requires a repulsion parameter r that is used to control the tightness of the clusters.

Practical [2] is a parameter-free method which scans the data and assigns each basket to an existing cluster or to a new one. The guiding quality function is akin to the "tf-idf" score used in information retrieval. *Practical* moves the baskets from a cluster to another while maximizing this function. It is able to automatically identify the clusters in presence of rare items.

TX-Means [6] is a parameter-free algorithm which partitions transactional data and extracts representative itemsets from each cluster. *TX-Means* is a top-down approach: it starts from a unique set of transactions and then splits them into partitions. The splitting principle follows the Bayesian Information Criteria (BIC) [15] of the original cluster versus the two subclusters. *TX-Means* uses the representative transactions to calculate the BIC. *TX-Means* peculiarities are *(i)* the usage of a local criterion instead of the maximization of a global function, and *(ii)* the usage of a centroid-based approach instead of a scanning schema.

4.3 On the Methods Equivalence

Here we look at the six methods and we classify them according to the following dimensions: design, optimization, order, and features.

Design. This dimension refers to the way an algorithm explores the result space. A *bottom-up* approach starts from all observations as independent clusters and looks for the best way to aggregate them. A *top-down* approach starts with all observations clustered in the same group and looks for the best splits. *TX-Means* is the only top-down approach in this paper.

Optimization. This dimension rules the way the splits and merges are performed. We distinguish between *global* and *local* criteria. A global criterion takes the relationship between the entire dataset and the partition into account (*Louvain*, *Infomap*, *Practical*, *Clope*), while a local criterion will only consider what is only directly connected to the each entity (*LabProp*, *TX-Means*). Global criteria can be further divided in several subclasses, of which we consider two. *Practical* and *Infomap* score differently nodes/items depending whether they are common or rare. *Clope* and *Louvain* score weights considering the ratio between the items/nodes in a cluster with respect to those outside the community/cluster.

Order. An *order-independent* method is deterministic: each run of the algorithm will always return the same clusters. An *order-dependent* method will adopt a randomized greedy strategy returning similar, but different, clusters across different runs. Here, *TX-Means* is the only order-independent approach.

Features. Here we look at differences in input/output of the various methods. Considering the input, all the methods are *parameter-free* with the exception of

Clope, which requires the repulsion parameter r. We set $r = 1.5$ by default, as it empirically shows good performances on all the different datasets analyzed. As for the output, *Infomap* and *TX-Means* return a hierarchical structure. *TX-Means* also returns a representative for each community/cluster.

5 Experiments

In this section we solve the community discovery problem $\mathcal{CD}$ on real datasets with ground truth. We want to compare the performance differences between community discovery methods and transactional clustering algorithms.

5.1 Datasets

Our data comes from SNAP[2]. We chose three *small* and three *large* networks. Table 1 reports a description of the networks and their ground truth communities.

Table 1. Networks and communities descriptions.

Network	Nodes	Edges	Density	Avg. clus. coef.	Communities	Avg. com. size
`karate`	34	78	0.139	0.571	2	17.0
`dolphins`	62	159	0.084	0.259	2	31.0
`football`	115	613	0.094	0.403	12	9.58
`amazon`	16,716	48,739	3.5e−4	0.649	5,000	13.49
`dblp`	93,432	335,520	7.7e−5	0.715	5,000	22.45
`youtube`	39,841	224,235	2.8e−4	0.197	5,000	14.59

The `karate` network is the social network between members of a karate club at a US university. The `dolphins` network is a network of frequent associations between dolphins living off New Zealand. The `football` network is the network of American football games between Division IA colleges during Fall 2000. In the `amazon` network we have a link between frequently co-purchased products. The Amazon product category defines the ground-truth community. The `dblp` network is a co-authorship network where two authors are connected if they published at least one paper together. Authors who published to a certain journal/conference form a community. Finally, in the `youtube` social network users form friendship with each other and can create groups. The user-defined groups are ground-truth communities. For these three datasets we consider the top 5,000 communities with highest quality [21]. Nodes which are not part of any of these 5,000 communities are dropped from the network.

[2] https://snap.stanford.edu/data/#communities.

5.2 Evaluation Measures

To evaluate the clustering quality we quantify the similarity between clusters and ground truth with the *Normalized Mutual Information (NMI)* [19]. *NMI* is preferred over *purity* because *(i)* it is more sensitive to the change in the clustering results, and *(ii)* it takes into account unbalanced distributions and does not necessarily improve when the number of clusters increases (as purity does). Given two sets of clusters $\mathcal{C}=\{c_1 \dots c_k\}$ and $\mathcal{G}=\{g_1 \dots g_{k'}\}$: $\text{NMI}(\mathcal{C},\mathcal{G}) = \frac{I(\mathcal{C},\mathcal{G})}{0.5*H(\mathcal{C})+0.5*H(\mathcal{G})} \in [0,1]$, where $I(\mathcal{C},\mathcal{G})=\sum_k \sum_j \left(\frac{|c_k \cap g_j|}{N}\right) \log \left(\frac{N|c_k \cap g_j|}{|c_k||g_j|}\right)$ is the mutual information [19], and $H(\mathcal{C}) = -\sum_k \left(\frac{|c_k|}{N}\right) \log \left(\frac{|c_k|}{N}\right)$ is the Shannon entropy [16]. Aligned partitions have a *NMI* $\sim$1, misaligned partitions $\sim$0.

We also evaluate the *deviation* $\delta_k=|\mathcal{C}|-|\mathcal{G}|$ between the real number of clusters and the number of clusters detected: $\delta_k \sim 0$ when the right number of cluster is detected, $\delta_k > 0$ if more clusters than real ones are detected, $\delta_k < 0$ otherwise.

5.3 Results

We report in Table 2 the results of the experiments performed by solving the community discovery problem both for community discovery algorithms and also for transactional clustering algorithms, after having transformed the network according to the procedure described in Sect. 3.

Table 2. Experiments evaluation through NMI and δ_k. Best performer for each dataset and each measure highlighted in bold. Ties are not broken.

Network	Infomap		Louvain		LabProp		TX-Means		Clope		Practical	
	NMI	δ_k	NMI	δ_k	NMI	δ_k	NMI	δ_k	NMI	δ_k	NMI	δ_k
`karate`	0.71	**1**	0.27	3	0.23	2	0.73	2	0.43	3	**0.77**	2
`dolphins`	0.57	4	0.04	4	0.01	**2**	0.53	6	0.52	10	**0.77**	**2**
`football`	**0.94**	**0**	0.20	−1	0.22	−1	0.87	−2	0.91	3	0.88	8
`amazon`	**0.87**	−3,893	0.50	−3,877	0.54	−3,375	0.68	**370**	0.84	−1,583	0.68	−2,162
`dblp`	0.64	−4,240	0.59	−4,417	0.75	4,239	0.60	**−563**	0.86	−1,390	**0.88**	−1,233
`youtube`	0.39	−4,349	0.27	−4,037	0.32	−3,756	0.38	**−408**	**0.72**	12,024	0.33	3,654

First of all, we underline the good performance of transactional clustering algorithms. In particular, with respect to the *deviation* δ_k it seems that clustering algorithms are better than community discovery methods in approximating the right number of communities for large networks.

Secondly, we observe that *Clope* is resulting to be one of the best performer in general in terms of NMI, but in the case of the `youtube` dataset it overestimates the number of communities, producing more than the double of the number of the real communities. As already discussed in [6], *TX-Means* is the best algorithm in keeping low the *deviation* δ_k, at least in larger networks.

Finally, we notice how the clustering algorithms are able to maintain good NMI performance both for small and large dataset with very different levels of density, nodes overlapping and average community size while some community discovery methods have very poor performance in the small networks. In summary, there is not clear comparative advantages in community discovery for community discovery methods, and more often than not a transactional clustering algorithm will be the best performer.

To conclude the section, we depict in Fig. 2 the results of a community discovery algorithm (Infomap) and a transactional clustering one (*TX-Means*). We can see that the results are quantitatively the same, as the NMI values are almost equivalent. However, *Infomap* and *TX-Means* have different points of failure, which might make one more suitable in some scenarios than the other.

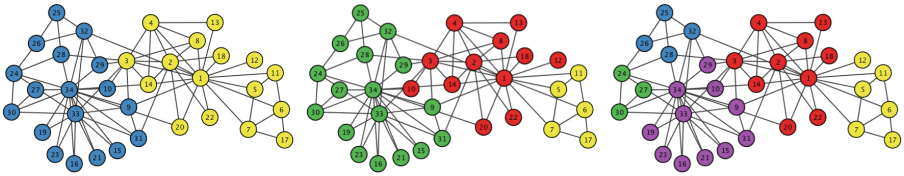


Fig. 2. The `karate` results: ground truth on the left, community discovery algorithm (*Infomap*, center), transactional clustering algorithm (*TX-Means*, right).

TX-Means returns a higher number of communities than the ground truth. However, it perfectly characterizes the boundary between the ground truth communities, and the additional ones are simply sub communities, hierarchically contained in the main two. On the other hand, *Infomap* is closer in number of communities, by returning only one of the three sub communities. However, its community boundary detection is worse than *TX-Means*: node 10 should have been classified on the leftmost community instead of in the rightmost one.

Finally, in Fig. 3 we highlight another common characteristic of *TX-Means* and *Infomap*: their ability to enhance which are the *core* nodes of each community. *Infomap* considers as core nodes those with the highest page rank score, while for *TX-Means* we assume the core nodes are those in the representative

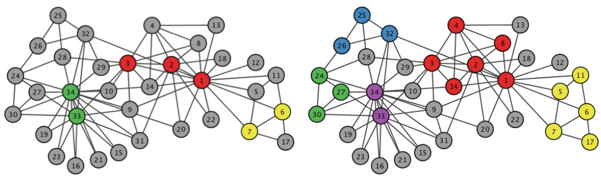


Fig. 3. The `karate` communities core nodes using a community discovery algorithm (*Infomap*, left) and a transactional clustering algorithm (*TX-Means*, right).

transactions. Figure 3 shows how *TX-Means* (right) is more "conservative" than *Infomap* (left): *TX-Means* mainly returns cliques of nodes delineating the backbone of the communities, *Infomap* mainly returns the hubs of the communities.

6 Conclusion

In this paper we have considered the problems of community discovery and of clustering transactional data and we have provided a formal mechanism to map the former into the latter. We have illustrated the main methodological similarities and differences among existing community discovery and transactional clustering approaches. Through some experiments, we have showed that such transformation empowers transactional clustering algorithms to perform as well as, or sometimes better than, specialized community discovery algorithms.

This paper can be the starting point for several interesting research directions. A very promising one is the extension of our mapping to show that community discovery $\leftrightarrow$ clustering. To add the left equivalence, we need a methodology to clean noise from transactional data. One promising approach to do it is to use algorithms developed for solving the network backboning problem [4].

Acknowledgements. This work is partially supported by the European Project *SoBigData: Social Mining & Big Data Ecosystem*, http://www.sobigdata.eu, GS501100001809, 654024. Michele Coscia has been partly supported by FNRS, grant #24927961.

References

1. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. JSTAT **2008**(10), P10008 (2008)
2. Bouguessa, M.: A practical approach for clustering transaction data. In: Perner, P. (ed.) MLDM 2011. LNCS (LNAI), vol. 6871, pp. 265–279. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23199-5_20
3. Coscia, M., Giannotti, F., Pedreschi, D.: A classification for community discovery methods in complex networks. SADM **4**(5), 512–546 (2011)
4. Coscia, M., Neffke, F.M.: Network backboning with noisy data. In: ICDE, pp. 425–436. IEEE (2017)
5. Coscia, M., Rossetti, G., Giannotti, F., Pedreschi, D.: Uncovering hierarchical and overlapping communities with a local-first approach. TKDD **9**(1), 6 (2014)
6. Guidotti, R., Monreale, A., Nanni, M., Giannotti, F., Pedreschi, D.: Clustering individual transactional data for masses of users. In: KDD (2017)
7. Huang, X., Lai, W.: Clustering graphs for visualization via node similarities. J. Vis. Lang. Comput. **17**(3), 225–253 (2006)
8. Li, H., Nie, Z., Lee, W.-C., Giles, L., Wen, J.-R.: Scalable community discovery on textual data with relations. In: CIKM, pp. 1203–1212. ACM (2008)
9. Malliaros, F.D., Vazirgiannis, M.: Clustering and community detection in directed networks: a survey. Phys. Rep. **533**(4), 95–142 (2013)
10. Niu, S., Wang, D., Feng, S., Yu, G.: An improved spectral clustering algorithm for community discovery. In: HIS, vol. 3, pp. 262–267. IEEE (2009)

11. Peel, L., Larremore, D.B., Clauset, A.: The ground truth about metadata and community detection in networks. Sci. Adv. **3**(5), e1602548 (2017)
12. Raghavan, U.N., Albert, R., Kumara, S.: Near linear time algorithm to detect community structures in large-scale networks. Phys. Rev. E **76**(3), 036106 (2007)
13. Rosvall, M., Bergstrom, C.T.: Maps of random walks on complex networks reveal community structure. PNAS **105**(4), 1118–1123 (2008)
14. Satuluri, V., Parthasarathy, S.: Scalable graph clustering using stochastic flows: applications to community discovery. In: KDD, pp. 737–746. ACM (2009)
15. Schwarz, G., et al.: Estimating the dimension of a model. Ann. Stat. **6**(2), 461–464 (1978)
16. Shannon, C.: A mathematical theory of communication. SIGMOBILE **5**, 3–55 (2001)
17. Tan, P.-N., Steinbach, M., Kumar, V., et al.: Introduction to Data Mining, vol. 1. Pearson Addison Wesley, Boston (2006)
18. Tsuda, K., Kudo, T.: Clustering graphs by weighted substructure mining. In: ICML, pp. 953–960. ACM (2006)
19. Vinh, N.X., et al.: Information theoretic measures for clusterings comparison: is a correction for chance necessary? In: ICML, pp. 1073–1080. ACM (2009)
20. Wang, F., Li, T., Wang, X., Zhu, S., Ding, C.: Community discovery using non-negative matrix factorization. DAMI **22**(3), 493–521 (2011)
21. Yang, J., Leskovec, J.: Defining and evaluating network communities based on ground-truth. Knowl. Inf. Syst. **42**(1), 181–213 (2015)
22. Yang, Y., Guan, X., You, J.: Clope: a fast and effective clustering algorithm for transactional data. In: TKDE, pp. 682–687. ACM (2002)

Recognizing Residents and Tourists with Retail Data Using Shopping Profiles

Riccardo Guidotti(✉) and Lorenzo Gabrielli

KDDLab, ISTI-CNR, Pisa, Italy
{riccardo.guidotti,lorenzo.gabrielli}@isti.cnr.it

Abstract. The huge quantity of personal data stored by service providers registering customers daily life enables the analysis of individual fingerprints characterizing the customers' behavioral profiles. We propose a methodological framework for recognizing residents, tourists and occasional shoppers among the customers of a retail market chain. We employ our recognition framework on a real massive dataset containing the shopping transactions of more than one million of customers, and we identify representative temporal shopping profiles for residents, tourists and occasional customers. Our experiments show that even though residents are about 33% of the customers they are responsible for more than 90% of the expenditure. We statistically validate the number of residents and tourists with national official statistics enabling in this way the adoption of our recognition framework for the development of novel services and analysis.

Keywords: Residents tourists classification
Customer shopping profile · Retail data · Spatio-temporal analytics
Data mining

1 Introduction

The availability of huge amount of personal data stimulates challenging questions that can be answered with data mining methodologies. Given a service providing spatio-temporal information, a question addressed in the last years for enabling the development of analysis and applications for social good pursues the goal of *recognizing* if a user of can be categorized as a *resident* or as a *tourist*. Since ground truth categories are generally missing, the resident-tourist classification is a very hard problem because supervised learning algorithms can not be applied. Instead, unsupervised approaches recognizing typical aspects of residents and visitors are used. Indeed, a number of previous works propose unsupervised methods by analyzing phone call data [5,7,12] or social media [4,10].

In this paper we propose an unsupervised methodological *framework* able to *recognize* residents and tourist by exploiting retail shopping data. Retail data is a very complex type of data containing various dimensions: *what* customers buy,

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 353–363, 2018.
https://doi.org/10.1007/978-3-319-76111-4_35

when and *where* they make the purchases and which is the *amount* of the purchase. Most of the works in the literature focus on *what* customers buy [1], while just a few of them exploit the *spatio-temporal dimension* [9,15]. To the development of our *recognition framework* we define a *temporal purchasing profile* capturing the shopping habit of a customer in terms of *when* and *where* she purchases and *how much* she spends. Then, we develop *data-driven* heuristic rules for categorizing the temporal shopping profile of each customer as `resident`, `tourist`, `occasional` and `rare`. In our vision `resident` customers live in a place close to their favorite store and show a continuous, uniform and considerable presence over all the monitored periods; `tourist` customers are people staying in a place close to a store only for a limited time but repeatedly along different periods; `occasional` customers purchase on a certain store in a not very frequent way but uniformly over all the monitored periods; `rare` customers performed only a single purchase on the monitored periods.

We instantiate our recognition framework for a case study on real retail transactions. The dataset analyzed was provided by *UniCoop Tirreno* which serves more than a million of customers on the west coast of central Italy. The dataset contains retail market data from January 2007 to December 2015. After customers categorization as `resident`, `tourist`, `occasional` and `rare`, we perform a range of exploratory analysis. Our main findings reveal that `residents` are responsible for more than the 90% of the expenses even though they are only the 33% the customers population, while the presences and expenses of `tourists` become significant especially in summer months and on the stores situated in vacation places, while in the other periods and places the largest part of the revenues is generated yet by `residents`. `Occasional` customers desultorily purchases at UniCoop stores but constitute 30% of the active customers. Moreover, we statistically validate our estimations of residents and tourists by performing correlation analysis and comparing our trends against official national statistics.

Summarizing, this work advances the achievements of existing works (Sect. 2) by *(i)* formalizing and generalizing the concepts of *temporal purchasing units* and *temporal shopping profile*, *(ii)* defining data-driven heuristic rules to *categorize customers* as residents, tourists or others (Sect. 3), *(iii)* showing an instantiation of the framework on a real case study for retail data which reports interesting findings, and proves reliable quantification with respect to official statistics (Sect. 4). Finally, we illustrate which are the applications that could exploit our recognition framework, and we outline future research directions (Sect. 5).

2 Related Work

Residents and tourists recognition has been accomplished for different purposes in different data domains. In the literature there exists a large set of works classifying users as resident and tourists by using social media and phone call data.

In [4], the authors separate users in residents and visitors in order to *study migration patterns* and to analyze the spread of the influenza like illness infection by monitoring Twitter posts. The approach developed in [10] for *investigating global mobility patterns*, assigns as the country of residence of a user the one in which she tweeted the most, while she is a visitor in all the other countries.

Other approaches, like the one presented in this paper, make the separation with the purpose of *analyzing the different characteristics* between the locals and the visitors. In order to provide space-time visual analytics of where the Seattle locals tweet and what they talk about, in [2] the authors profiles a user as local or visitor, by counting the days in which a user tweeted inside or outside Seattle. Similarly, to mine the mobility patterns of tourists in Florence, in [8] all the users who posted geolocated Flickr photos for less than 30 days in the province are considered tourists. In [14] the users are classified as resident and tourists by analyzing the number of user's active days in a specific area.

However, the methods of discrimination between residents and tourists described above are nothing more than a set of more or less sophisticated rules that are primarily dictated by common sense rather than by a data-driven finding. In addition, a simple rule-based approach to partition locals and visitors may hinder to derive some more useful information obtainable by conducting a deeper analysis on the data to derive the behavior of the users.

To overcome the aforementioned limitations, the authors of [5,6] defined how to build individual profiles based on mobile phone calls such that the profiles are able to characterize the calling behavior of a user. By analyzing these profiles three categories of users are identified: *residents*, *commuters* and *visitors*. In [7] this characterization is strengthen by aggregating users having a similar calling behavior with the *k-means* clustering algorithm [19]. The centroid of each cluster is compared with pre-defined prototypes representing the categories of interest, then, each cluster is classified by means of the associated prototype. The proposed framework aims at defining residents and tourists profiles by outlining the data-driven methodology identifying the users habits described in [5–7,12]. In [13] is proposed an alternative unsupervised procedure for estimating the tourists presence in an area over a specific period of time.

Finally, it is worth to underline that residents-tourists partitioning is useful for a myriad of studies ranging from economy to sociology to mobility. In [11] it is observed the economic impact of special events in locals and visitors. In [3] it is analyzed the relationship between place attachment and landscape value studying subgroups of resident and visitors having different levels of reliability with respect to the measures considered in the analysis. Moreover, interesting understanding of tourism lifecycle models comes from [16,18] where the impact that locals and tourists have on museums and shopping spaces are investigated.

3 Residents and Tourists Recognition

In this section we define the analytical framework able to classify the customers of a retail market chain into data-driven behavioral categories.

The type of data analyzed by the framework consists in customer's shopping sessions. A *shopping session* contains information about *(i)* which customer made the purchase, *(ii)* all single items composing the basket, *(iii)* in which shop the transaction happened, *(iv)* the time and the *date* of the shopping session. We underline that in every retail market chain the purchases can be individually assigned to a customer only if a fidelity card is used for the purchases.

In order to recognize residents and tourists among the customers, the framework considers only the information relative to the *date* of the shopping sessions. The other information are exploited to characterize the categories recognized.

Customer Shopping Model. For each customer, we summarize a set of shopping sessions by introducing the *temporal purchasing unit* (unit in short):

Definition 1 (Temporal Purchasing Unit). *Given a period τ, we define the* temporal purchasing unit *u of customer c as a vector $u \in \mathbb{R}^M$, where M is the number of time-parts considered, and u_i contains the number of purchases performed by customer c in the i-th time-part.*

With *time-part* we indicate any aggregation of days, e.g. day, weeks, months, etc. Given a period τ, for every customer we can observe a set $U_c = \{u^{(1)}, \ldots, u^{(N)}\}$ of temporal purchase units. For example, in Fig. 1 are shown three temporal purchasing units for a customer where as period τ is used the year and as time-part the month, i.e., $M = 12$. Every line represents the number of purchases: the higher the line, the higher the number of purchases in a month.

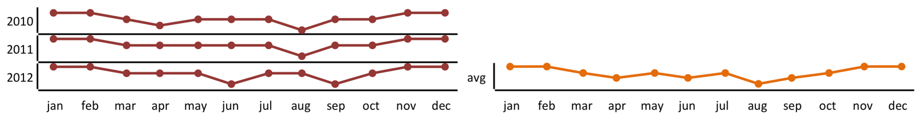


Fig. 1. *Left:* temporal purchasing units. *Right:* temporal purchasing profile. The higher the line the more purchases are done in that month.

We define the *temporal purchasing profile* (profile in short) of a customer as:

Definition 2 (Temporal Purchasing Profile). *Given the set of temporal purchase units $U_c = \{u^{(1)}, \ldots, u^{(N)}\}$ of customer c, we define the* temporal purchasing profile *$p^{(c)}$ of c as an aggregation of U_c into a vector $p^{(c)} \in \mathbb{R}^M$, where M is the number of time-parts considered, and $p_i^{(c)} = \frac{1}{N} \sum_{u^{(j)} \in U_c} u_i^{(j)}\ \forall\, i = 1, \ldots, M$ contains the average number of purchases performed by c in the i-th time-part.*

The approach for the extraction of the profiles can be summarized as follows. The *first* step is the *construction* of the temporal purchase units from the shopping sessions considering M time-parts for each period τ. Then, the *second* step consists in aggregating the temporal purchase units $U^{(c)}$ of each customer to obtain the temporal purchasing profiles $p^{(c)}$ according to the previous definition.

Unsupervised Customers Classification. Our aim is to find prototypes describing the behavior for `residents`, `tourists`, `occasional`, and `rare`.

As preprocessing step, we label as `rare` all the customers that have performed only one purchase in the purchasing profile $p^{(c)}$, i.e., such that $\sum_i^M p_i^{(c)} = 1$.

Given a set of non rare customers $C = \{c_1, \ldots, c_N\}$ each one with her temporal purchasing profile $P = \{p^{(c_1)}, \ldots, p^{(c_N)}\}$, our objective is to find a partitioning $\mathcal{G} = \{G_1, \ldots, G_K\}$ of P into K disjoint sets. In other words, we want to cluster the customers in C according to their profile such that customers having a similar temporal shopping behavior belong to the same cluster.

To perform such partitioning we employ the well known *k-means* clustering algorithm [19]. Since in real applications the temporal purchasing profile is in fact a not sparse vector with a treatable size, we run k-means by employing the *euclidean distance* [19] as distance function between two profiles.

From the partitioning we obtain K clusters $\mathcal{G}=\{G_1, \ldots, G_K\}$ of similar customers, and K centroids $g^{(1)}, \ldots, g^{(K)}$ calculated by computing the mean for each cluster, i.e., $g_i^{(k)}=\frac{1}{|G_k|}\sum_{p^{(c)} \in G_i} p_i^{(c)}\ \forall\, i = 1, \ldots, M,\ \forall\, k = 1, \ldots, K$.

The last step consists on assigning for each *not rare* customer $c_i \in C$ a behavioral label by observing the centroid $g^{(k)}$ corresponding to the cluster G_k the customer c_i belongs to. We indicate the centroid $g^{(k)}$ of customer c with the notation $g^{(c)}$, i.e., $p^{(c)} \in G_k$. After having empirically observed various different shapes and values for the trends described by the centroids $\{g^{(1)}, \ldots, g^{(K)}\}$ we identified three categories: `resident`, `tourist` and `occasional` (see examples in Sect. 4), and we developed the following heuristic to assign the label:

- if any peak is detected in $g^{(c)}$ we label c as `tourist`, otherwise;
- if $\frac{1}{M}\sum_i^M g_i^{(c)} \leq 0.5$ we label c as `occasional`, otherwise;
- $g^{(c)}$ shows a considerable and uniform trend and we label c as `resident`.

In practice, we classify as `tourist` the customers appearing only in a precise moment with respect to the time period τ and number of time-parts M, as `occasional` those with a very low average number of purchases, and as `resident` the customers with a shopping trend which is almost constant and not negligible.

4 Case Study

In this section we present a case study on a massive real dataset of customer's shopping sessions showing the effectiveness of the proposed approach in terms of quantification and qualification of the category recognized.

Dataset. We develop our case study on the *Coop* dataset provided by *UniCoop Tirreno*[1], one of the largest Italian retail distribution company. It serves more than a million of customers covering an extensive part of the Italian territory. The stores mainly cover the west coast of central Italy. The shop distribution is not homogeneous (see Fig. 2): shops are located in a few Italian regions (Tuscany,

[1] https://www.unicooptirreno.it/.

Lazio, Campania) and therefore, the coverage of these regions is much more significant, while customers from other regions usually shop only during vacation periods in these regions. The 138 stores sell about 8,000 different items.

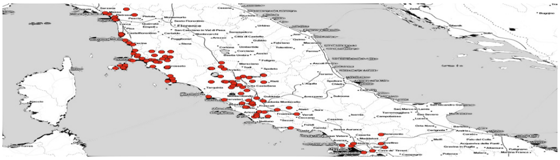


Fig. 2. Coop stores positioning on Italy west cost.

The dataset contains retail data from January 2007 to December 2015 belonging to 1,637,311 active and recognizable customers. A customer is active if she purchased during the time window, while she is recognizable if the purchase has been made using the membership card which can enable to discounts. The company is able to tie each shopping session to the card, and, for each shopping session the company knows all the information described in Sect. 3.

Customers Categorization. On top of the *Coop* dataset we instantiate the framework described in Sect. 3. First, we build the temporal purchase units from the shopping sessions with $M = 12 \times 2 = 24$ time-parts, i.e., a time-part per month considering separately weekend and weekdays, and a purchase unit for each period where $\tau = 1$ year. Given a customer c, the i-th element of the vector $u^{(j)}$ contains the number of purchases performed by c in a certain month-weekend/weekday for a certain year. Consequently, the i-th element of the temporal purchasing profiles $p^{(c)}$ contains the average number of purchases performed by c in a certain month-weekend/weekday with respect to the years c was active. Then we label as `rare` the 306,091 customers having only one purchase, and we run the *k-means* algorithm on the profiles of the remaining customers. We set the number of clusters $K = 20$ in correspondence of the position of the knee in the SSE curve [19]. Finally, we label the centroids (and the customers belonging to the clusters) according to the categorization rules defined in Sect. 3. As consequence, we highlight that small variation in the choice of K would still lead to the behaviors described in the following and delineating the four different data-driven groups (`rare`, `resident`, `tourists`, `occasional`).

We report a sample of the centroids $g^{(1)}, \ldots, g^{(K)}$ in Fig. 3. In the first row we observe a behavior typical of `resident` customers. They purchase constantly and uniformly along the whole year with a different level of repetitiveness: 3-4 times per weekday-month for the customers in *cluster 06*, 1-2 times per weekday-month for the customers in *cluster 04*. In the second row are reported centroids of customers labeled as `tourists`. They become active only in specific months: June, July, August, and they are also markedly active also on weekends. It is worth to underline that most of these tourists are in fact customers (with a

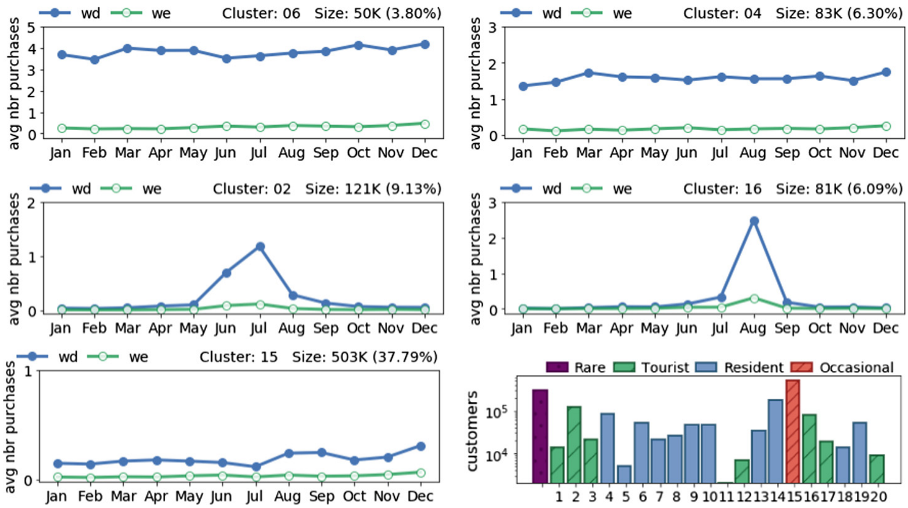


Fig. 3. Sample of clusters centroids. *First* row centroids for `residents`, *second* row centroids for `tourists`, *third* row centroid for `occasional` and clusters' sizes.

fidelity card) repeatedly appearing along different years. They are not "one-time tourists". Finally, the third row shows the centroid for `occasional` customers.

Table 1. Quantification of customers, expenditure, purchases and number of clusters.

	#customers	%customers	%expenditure	%purchases	#clusters
`residents`	554,243	33.86	91.76	93.65	11
`tourist`	273,934	16.73	02.97	02.57	8
`occasional`	503,043	30.72	05.11	03.63	1
`rare`	306,091	18.69	00.19	00.15	1

Details about the categorization are reported in Table 1. According to the Pareto principle [17], more than 90% of expenses and purchases are generated by 33% of the customers which are labeled as `residents`. Moreover, we observe that `tourists` are half of the `occasional` and generate almost the same level of expenditure and number of purchases. We remark that *all* the customers belonging to every category have a fidelity card that could have been subscribed not necessarily with *UniCoop Tirreno* but in any other *Coop* store across Italy (e.g. UniCoop Firenze, Coop Lombardia, etc.), allowing in this way every Coop (including *UniCoop Tirreno*) to trace the purchases made with the fidelity card.

Figure 4 shows how change the presences *(left)* and the expenditure *(right)* of the customers for the various categories along different months. In *summer months* we observe an increasing percentage of both `tourists` and `rare` customers, while the proportion of `residents` and `occasional` remains constant.

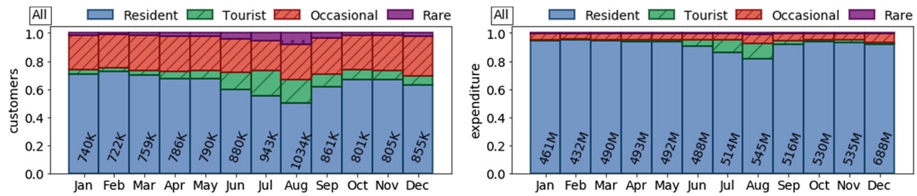


Fig. 4. Stacked bar plot with respect to relative number of customers *(left)* and relative expenditure *(right)*. The bottom numbers reports the absolute sum of the values.

This phenomenon is not due to a decrease of these last two categories, but rather to an increment on the overall number of active customers how highlighted by the absolute values on the bottom. In line with the statistics above, the strong increment in the proportion of the presences of `tourists` and `rare` customers causes only a small increment in the proportion of the expenditure. In Fig. 4 we also notice how even though `tourists` are much less than `occasional` along the year, in July and August `tourists` expends more than twice as much as `occasional` do, confirming that `tourists` are not just "occasional" sightseers but customers repeatedly following this specific behavior. Finally, it is worth to underline that August is the month with the highest number of presences but December is the month with the highest expenditure, while August is only second in terms of expenditure. Together with the categorization this indicates that in December `residents` markedly increase their expenditure while in August the expenditure increase is due to `tourists`, `occasional` and `rare` customers.

Validation of Category Quantification. We validate our estimations of `tourists` and `residents` by comparing the quantification returned by our framework using the *Coop* dataset with the official Italian data provided by *ISTAT*[2]. We remark that we are estimating if the *quantification* resulting from the unsupervised categorization is in line with the official statistics: we can not validate the *classification* because we do not have the ground truth for the customers. Moreover, not comparable methods to use as competitors are present in the state of the art. We employ two different indicators to estimate the correlations [19]. *Pearson* evaluates the linear relationship. A relationship is linear when a change in one variable is associated with a proportional change in the other variable. *Spearman* evaluates the monotonic relationship. In a monotonic relationship, the variables tend to change together, but not necessarily at a constant rate.

We validate the estimation of tourists by observing the trends along different months and years. Since ISTAT does not provide the tourists for each month at regional level, we compare our estimation with the national trend (*all-istat*). How showed on Fig. 5 *(left)*, *all-istat* and the sum of the trends of Tuscany, Lazio and Campania *west-istat* are identical (Pearson and Spearman of 0.99). Figure 5 *(center)* shows a very high similarity between the `tourist`-*coop* and *all-*

[2] http://dati.istat.it/, http://demo.istat.it/.

istat tourists trends on a monthly granularity from 2010 to 2015[3]. These trends reports a Pearson of 0.91 (p-value 7.95e−29) and a Spearman of 0.90 (p-value 4.59e−28). As consequence, we can claim that *(i)* the trend of tourists quantified by the framework is a good proxy for official statistics, *(ii)* the customers in the `tourists` cluster follow the real trend of tourists and consequently they can be effectively real tourist-customers which systematically reappear across the years. Moreover, we observe in Fig. 5 (right) the `rare`-*coop* and the *all-istat* tourists trends with a Person of 0.95 (p-value 1.81e−37) and a Spearman of 0.92 (p-value 3.75e−30). This means that `rare` customers are "one-time tourists" and also in this case they are a good approximation for nowcasting official statistics.

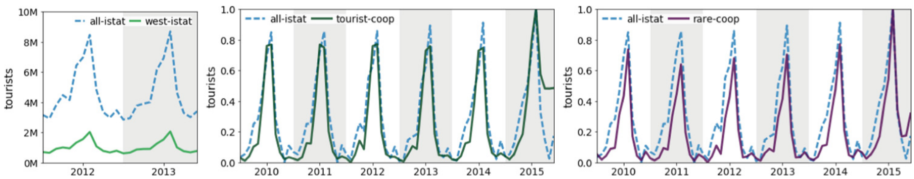


Fig. 5. Trends comparisons. ISTAT national (*all-istat*) against: (*left*) ISTAT sum of tourists in west cost (*west-istat*), (*center*) *tourist-coop*, and (*right*) *rare-coop*.

We validate the `residents` by observing the correlations with respect to different municipalities. We observe an average Pearson of 0.68 (p-value 3.97e−47) and Spearman of 0.74 (p-value 3.24e−34). For these correlations we must consider that they are tied with the adoption of *UniCoop Tirreno* stores with respect to the size of the cities and with the presence of other supermarket chains.

Single Stores Analytics. The categorization returned by the framework together with the validation obtained in the previous section, empower detailed studies of some aspects of these sets of customers from different point of views. In particular, in this section we show how the analysis of single stores reveals a different type of "audience" with respect to the four categories.

Figure 6 reports an example of two very different stores in *Livorno* and *San-Vincenzo*. In *Livorno* (upper left) we observe a negligible `rare` customers and `tourists` along over the year and a consistent percentage of `occasional` customers, while with respect to the expenditure (upper right) there is a negligible `tourists` effect with the total number of customers and expenditure remaining stable along the whole year. On the other hand, on *San Vincenzo*, the number of customers in summer months is about three times the number of customers in the other months (lower left). In addition, about 40% of the total expenditure in July and August is caused only by `tourists` (lower right). The other stores present in the dataset can be similar to *Livorno* or to *San Vincenzo*, or have a customer audience which is "in the middle" with respect to them.

[3] Missing years are not available on the ISTAT website.

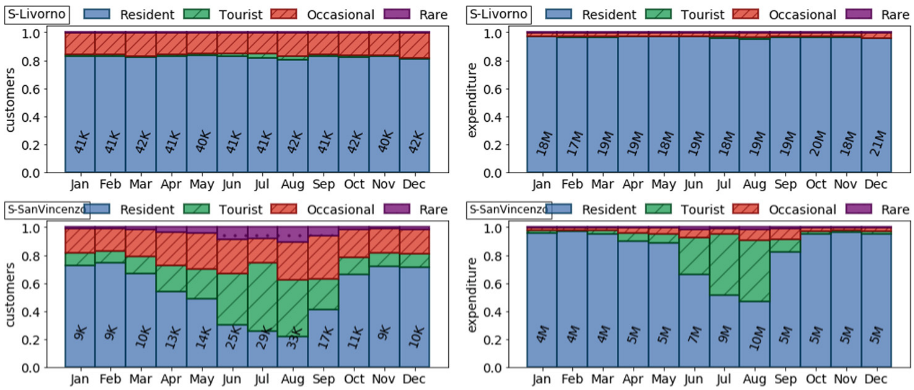


Fig. 6. Stacked bar plot with respect to relative number of customers *(left)* and relative expenditure *(right)* for two different stores: *Livorno* and *San Vincenzo*.

Therefore, the type of "audience" of a certain store can change or not along the year and, by exploiting the proposed categorization, store managers can build ad-hoc marketing strategies to maximize the profits in every month.

5 Conclusion

Nowadays various data sources, from mobile phone data to shopping transaction, are proxies for studying the user's social life. We have presented a framework for recognizing and estimating residents and tourists exploiting retail market data through the definition of temporal shopping profiles. We have applied our framework on a real massive dataset of transactions and we have observed that for Coop the residents are the customers belonging to the 20% of the Pareto principle. We have validated our estimation with national official statistics and on top of that we have showed that every store can be characterized with respect to the presences of tourists and residents and their expenditure in different months.

Our framework could be exploited in real applications for a common social good. The tourists and residents estimations can be adopted as *early approximations* of the official statistics which are provided long time after the period to which they relate. In addition, our framework enables peculiar estimations of the presences with a very fine grained spatio-temporal granularity. From the market chain perspective the customers groups can be used to better organize the product disposal and to plan targeted marketing campaigns, i.e., give to tourists special discounts on sun creams and summer products.

The framework proposed and the resulting customers categorization opens novel and interesting research directions. A very analytical step consists in better *characterizing* each group of customers to understand if they have particular preferences with respect to the time they go to shopping, to the products purchased, and to the most frequent patterns. Finally, the customers categorization together with information about the nationality can help in better defining both the *tourists flows* but also, with respect to the residents, the *migration flows*.

Acknowledgement. This work is partially supported by the European Project *SoBigData*, 654024, http://www.sobigdata.eu. We thank UniCoop Tirreno for allowing us to analyze the data and to publish the results.

References

1. Agrawal, R., et al.: Mining association rules between sets of items in large databases. In: ACM Sigmod Record, vol. 22, pp. 207–216. ACM (1993)
2. Andrienko, G., et al.: Thematic patterns in georeferenced tweets through space-time visual analytics. Comput. Sci. Eng. **15**(3), 72–82 (2013)
3. Brown, G., et al.: The relationship between place attachment and landscape values: toward mapping place attachment. Appl. Geogr. **27**(2), 89–111 (2007)
4. Cao, G., et al.: A scalable framework for spatiotemporal analysis of location-based social media data. Comput. Environ. Urban Syst. **51**, 70–82 (2015)
5. Furletti, B., Gabrielli, L., Renso, C., Rinzivillo, S.: Identifying users profiles from mobile calls habits. In: KDD Workshop, pp. 17–24. ACM (2012)
6. Furletti, B., Gabrielli, L., Renso, C., Rinzivillo, S.: Analysis of GSM calls data for understanding user mobility behavior. In: Big Data, pp. 550–555. IEEE (2013)
7. Gabrielli, L., Furletti, B., Trasarti, R., Giannotti, F., Pedreschi, D.: City users' classification with mobile phone data. In: Big Data, pp. 1007–1012. IEEE (2015)
8. Girardin, F., et al.: Understanding of tourist dynamics from explicitly disclosed location information. In: Symposium on LBS and Telecartography, vol. 58 (2007)
9. Guidotti, R., Coscia, M., Pedreschi, D., Pennacchioli, D.: Behavioral entropy and profitability in retail. In: DSAA, pp. 1–10. IEEE (2015)
10. Hawelka, B., et al.: Geo-located twitter as proxy for global mobility patterns. Cartogr. Geogr. Inf. Sci. **41**(3), 260–271 (2014)
11. Long, P.T., Perdue, R.R.: The economic impact of rural festivals and special events. J. Travel Res. **28**(4), 10–14 (1990)
12. Lulli, A., et al.: Improving population estimation from mobile calls: a clustering approach. In: ISCC, pp. 1097–1102. IEEE (2016)
13. Mamei, M., et al.: Analysis of tourist activity from cellular data. In: POSTER (2017)
14. Manca, M., et al.: Using social media to characterize urban mobility patterns: state-of-the-art survey and case-study. Online Soc. Netw. Media **1**, 56–69 (2017)
15. McDonald, W.J.: Time use in shopping: the role of personal characteristics. J. Retail. **70**(4), 345–365 (1994)
16. Moreno Gil, S., Ritchie, J.B.: Understanding the museum image formation process: a comparison of residents and tourists. JTR **47**(4), 480–493 (2009)
17. Pareto, V.: Manual of Political Economy. Macmillan, London (1971)
18. Snepenger, D.J., Murphy, L., OConnell, R., Gregg, E.: Tourists and residents use of a shopping space. Ann. Tour. Res. **30**(3), 567–580 (2003)
19. Tan, P.-N., et al.: Introduction to Data Mining. Pearson Education India, Noida (2006)

Advanced Interaction Technologies for Accessible and Engaging Cultural Heritage

Daniela Fogli[1(✉)], Davide Sansoni[1], Emanuele Trivella[1], Alberto Arenghi[2], and Ivana Passamani[2]

[1] Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Brescia, Brescia, Italy
daniela.fogli@unibs.it, davide.sansoni@gmail.com, emanuele.trivella@gmail.com

[2] Dipartimento di Ingegneria Civile, Architettura, Territorio e Ambiente e di Matematica, Università degli Studi di Brescia, Brescia, Italy
{alberto.arenghi,ivana.passamani}@unibs.it

Abstract. This paper describes a web-based system composed of an authoring tool and a cross-platform mobile application, based on augmented reality and Bluetooth Low Energy technology, aimed at improving the visitor experience in a museum through tailored, accessible and engaging content and interaction. The system proposes itself as a low-cost solution for museum organizations, both in terms of required technical devices (visitors' smartphone is exploited for experiencing the augmented reality solution) and professional skills needed for long-term content maintenance. As to the latter, the authoring tool allows museum curators to create and manage all necessary contents that make up an app instance, including the structure of the museum, artwork descriptions and related materials (photos, videos and speeches) suitable for the different user profiles. A preliminary experimentation of the system demonstrates the feasibility of the proposal.

Keywords: Augmented reality · Beacon · Content management system Universal design · Accessibility

1 Introduction

The Council of Europe defines Cultural Heritage (CH) as a "group of resources inherited from the past which people identify, independently of ownership, as a reflection and expression of their constantly evolving values, beliefs, knowledge and traditions" [7]. UNESCO, on the other hand, promotes both the preservation and enhancement of CH, two activities that can appear as opposite each other. Information and Communication Technologies (ICTs) are being proposed to address these needs, by fostering the enhancement and enjoyment of CH in new ways and, at the same time, by sensitizing people about its importance and fragility [19]. In particular, Augmented Reality (AR) offers opportunities to empower visitors' experience through the overlay of digital content onto objects, artifacts and environments (e.g. [12, 13]). AR perceived

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 364–373, 2018.
https://doi.org/10.1007/978-3-319-76111-4_36

value has been recently investigated from the perspective of different stakeholders [8]: it emerged that cultural organizations, and small museums in particular, often fear the costs of such new technologies and need to have suitable resources in terms of people with technical skills and of hardware/software technologies.

In the Framework Convention of the Value of CH for Society, the Council of Europe also underlines that "everyone, alone or collectively, has the right to benefit from the cultural heritage" [7]. This issue is often neglected in existing ICT solutions, which often present accessibility barriers and 'one-size-fits-all' contents; whilst, in the cultural sector, the content to be transmitted might not be the same for a child or for an expert, just like the communication media used for blind and deaf people might be different [2].

The UniBSArt4All project presented in this paper aims at addressing the issues mentioned above according to Universal Design principles. More precisely, this project has been developed around three main themes:

1. *User experience*. Enhancing user experience of cultural heritage through AR solutions based on automatic recognition of artwork and wireless devices.
2. *Accessibility*. Making both content and interaction accessible to different types of users.
3. *Sustainability*. Ensuring long-term sustainability of the solution, both financially and technically, through low-cost technology and a suitable authoring tool.

The project is the result of an interdisciplinary activity carried out by experts in ICT, in CH preservation and promotion, and in accessibility and universal design. In the frame of this project, we developed a cross-platform (Android and iOS) mobile application (*UniBSArt4All app* in the following), aimed to support users in enjoying museum content in a tailored and accessible way, and a content management system (*UniBSArt4All CMS* in the following), to be used by museum curators to populate the database of contents exploited by the app.

The system does not require visitors to interact with some invasive or external technology, such as head-mounted displays or smart glasses, but just with their personal smartphones. The solution is thus cheap both for visitors, who must only download a mobile app from the store, and for museum organizations, which must enrich rooms and objects with Bluetooth Low Energy (BLE) devices (beacons) for artwork identification, and possibly pay a service license for image recognition. Most of the effort required to museum curators consists of creating suitable contents for the different types of users that the application is able to support.

The whole system has been preliminarily experimented in a monumental complex belonging to the University of Brescia, in Italy. However, the idea is to provide museum curators with an integrated system supporting the development of tailored and advanced guides for museums.

The paper is organized as follows: Sect. 2 discusses related work about the adoption of AR in cultural heritage domains; Sect. 3 introduces the project UniBSArt4All by describing both the mobile application and the CMS; Sect. 4 briefly describes a preliminary experimentation of the app; while Sect. 5 concludes the paper.

2 Related Work

The adoption of AR is increasing in the CH domain, with specific reference to archeological sites and museums. Digital overlay of information on visitor's surrounding can be obtained using mobile or tablet devices [11], or smart glasses [20]. These solutions may provide for example 3D reconstructions of artworks and monuments [18, 21], or superimpose digital objects over real ones to allow accessing additional content. Otherwise, a video projector can be used to enrich the real world with digital information that several users may collaboratively enjoy [4]. Interaction techniques for exploring digital information could be moving the device from a layered menu [12], or rotating and shaking the device to send various types of commands. Head-mounted devices, usually adopted to provide Virtual Reality experiences, are an alternative solution to strive for immersion in a museum space with both seeing and gesturing: for instance, TombSeer exploits head-mounted displays to provide an AR experience and combine it with natural gestural interactivity [17]. This requires, however, that visitors wear an external device, provided by the museum, whose operation must be learned easily and whose cost should be kept low.

Our goal is instead to improve visitor's engagement with cultural heritage through an immersive interaction that is also easy to deploy and to accept by end users: using his/her personal smartphone the visitor can maintain the visual contact with the artwork through the photo camera, and access all additional contents about the artwork in a specific portion of the screen, as well as he/she can interact with these contents in different ways, by adjusting for example its font size or scrolling it or dealing with associated videos or speeches.

Interaction features are however the same provided by usual mobile applications and thus the user has not to learn specific gestures.

User profiling is often used in literature approaches to adapt the CH application to the users' preferences (e.g., [22]). In our system it is used for tailoring content and interaction experience according to possible user's disabilities or other characteristics (age and cultural background).

Another novelty with respect to previous solutions is the usage and experimentation of beacons, to allow the automatic detection of artworks, thus overcoming the problems that GPS encounters in indoor environment [11] and avoiding the need of approaching QR codes or fiducial markers in an environment potentially full of people.

Finally, solutions proposed in literature usually consist of ad-hoc systems designed for some specific CH site or museum; no indication about their development and maintenance over time, or deployment over different organizations is in generally provided. Our project started by taking into consideration both visitors' and curators' needs, in order to propose a feasible and financially viable solution.

3 UniBSArt4All

UniBSArt4All was iteratively developed starting from scenarios and mock-ups, and by taking constantly into account the three themes mentioned in the Introduction. AR technology was constantly regarded as a new medium of communication with CH and not as an aesthetic experience in itself; therefore, visitors, curators and human contexts remained at the center of the design activity [16].

This activity led to conceive the whole system architecture as shown in Fig. 1. In the following we focus on UniBSArt4All app, which proposes itself as a novel concept of tourist guide, beyond the audio guide [3]. Then, the structure and goals of the CMS will be briefly described.

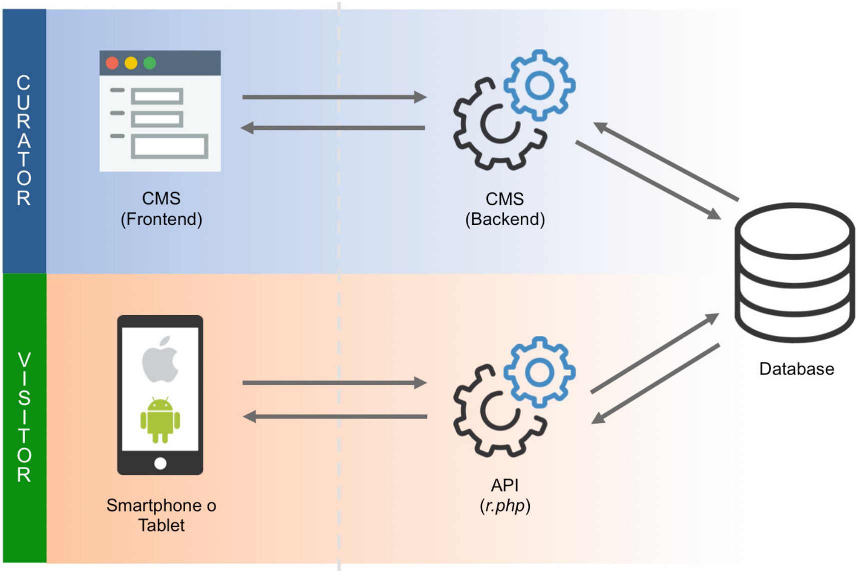


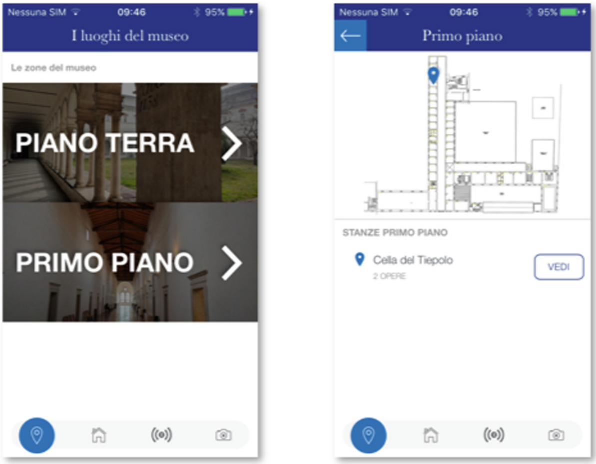
Fig. 1. The system UniBSArt4All.

3.1 UniBSArt4All App

The app was implemented through the Cordova framework, which allows cross-platform development by means of web-based languages. The app is structured along four sections:

- *Homepage* that provides first access to the app as well as the last museum news.
- *Museum*, which, according to a step-by-step interaction, shows the sequence of floors, rooms, artworks and artwork details (Fig. 2).
- *Bluetooth screen*, which uses beacon technology to detect the closest artwork and shows information related to that artwork (Fig. 3).
- *AR screen*, which uses the photocamera and the Wikitude service (https://www.wikitude.com) to recognize the artwork, and shows information related to the recognized artwork (Fig. 4).

The Bluetooth screen exploits a Cordova plugin that allows scanning the surrounding environment to look for beacons associated to artworks. Whenever a beacon is detected, the ID of the artwork is sent to the application and used to define a proper



(a) (b)

Fig. 2. Floors of the museum (a) and map related to the first floor with rooms and related artworks (b).

Fig. 3. Bluetooth screen.

query to the database for extracting all related contents that will be shown in the right part of the screen; the rest of the screen (2/3) will be reserved to the visualization of information captured by the camera to keep visual contact with the artwork (Fig. 3). Contents on the right might be scrolled and tailored to users' needs, e.g. fonts may be resized or full-screen can be activated. Whenever more than one beacon is detected, the application produces a vibration and presents the user with a popup including the list of the artworks related to the three closest devices. The list is ordered on the basis of the BLE distance from the visitor, who may then select the artwork he/she is interested in.

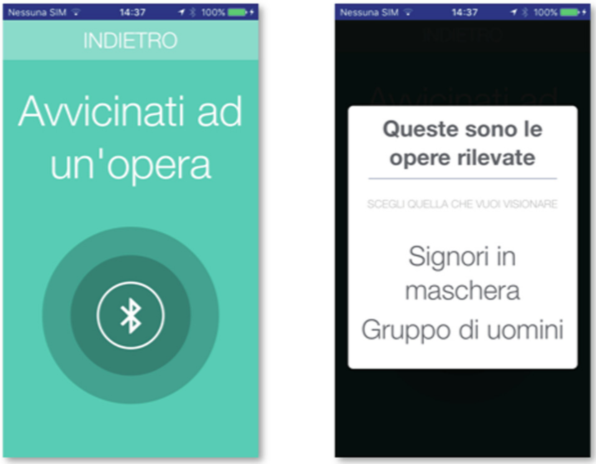
The AR screen has been implemented through the Wikitude framework: when the user enters in this screen, a Wikitude World is activated, i.e. a different application,

Fig. 4. AR screen based on artwork recognition with Wikitude service used in its trial version.

separated from the main application, performs artwork recognition by sending the image captured by the camera to the service made available in the Wikitude Cloud Recognition section. The latter includes a storage where the target images can be uploaded for subsequent comparison with the captured image. When artwork identification is successful, the same application extracts all information from the database and visualizes them in the right part of the screen, similarly to the Bluetooth screen.

The app allows tailoring content and interaction according to the accessibility needs of the user. In particular, the user may select a possible disability condition (hearing or visual impairment), thus allowing the app to adapt its interaction features. For instance, in the case of visual impairment, the app provides voice-over features and organizes the pages in a way suitable to support functionality selection by visually impaired users. In this case, the AR screen is disabled, while the Bluetooth screen provides the indication "Move close to an artwork" (Fig. 5a), which is automatically told by the screen reader available in the device. In case only one beacon is detected, a vibration is produced and then all contents about the artwork are told to the user. Otherwise, a simplified pop-up, suitable to visually impaired users, is shown (and told) to allow artwork selection (Fig. 5b). In case the user declares a hearing disability, proper contents will be loaded that propose videos, possibly enriched with subtitles or descriptions through sign language.

A second aspect of accessibility is related to the nature of contents, which should be adequate to the age, culture and objectives of the user. In particular, on first access to the app, one may declare him/herself as a child, a tourist or a scholar, and the app will select automatically the contents that are most suitable to that type of user. For instance, children will find artwork descriptions in a simplified language and will have the possibility to vote the artworks they like. The tourist profile is conceived for a user who is not expert in the cultural domain but would like to access more detailed information with respect to a child. Complete and detailed contents, using domain-oriented terminology, will instead be made available to scholars. These different types of content will be created and uploaded in the shared database by museum curators, by means of the CMS.



(a) (b)

Fig. 5. Interface for visually impaired visitors: artworks detected through beacons (a) and described by the app (b).

3.2 UniBSArt4All CMS

A web-based CMS has been specifically developed to support museum curators in the management of the app and its contents (Fig. 6). The CMS has been conceived to allow an easy adaptation of the project to any type of museum. The following sections compose it:

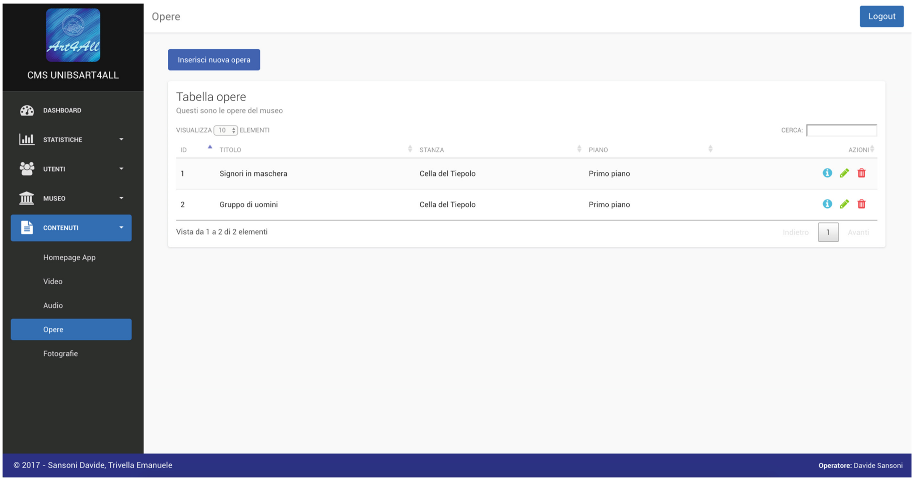


Fig. 6. UniBSArt4All CMS.

- *Dashboard*: it shows the visit trends in the current year classified according to user profiles.
- *Statistics*: it presents several graphics that describe the museum and artwork visits, classified on the basis of temporal intervals that the curator may personalize according to his/her needs.
- *Users*: this section allows monitoring user registrations to the app.
- *Museum*: this part allows curators to set up the basic data of the app, i.e. the museum floors and their related rooms, as well as to create artwork objects with automatic generation of IDs; furthermore it supports simple assignment of beacons to artworks.
- *Contents*: with this section the curators can take news updated, and upload and manage all contents related to artworks, such as photos, videos, textual and audio-descriptions. In particular, videos and descriptions must be properly created for the different user profiles.

In the current version of the system, artwork photos must be uploaded also in the Wikitude Cloud Storage, by inserting artwork IDs in the meta-data, in order to support their recognition through the Wikitude plugin.

4 Preliminary Experimentation

As already mentioned, UniBSArt4All was preliminary experimented in the San Faustino monumental complex of the University of Brescia, and in particular in the cloister of the monastery and in the Tiepolo cell, where one may admire some beautiful frescos. Three technical experimental sessions were carried out to test image recognition under different brightness, prospective and distance conditions, and to test the BLE technology. After some tuning of photo capturing, the app operated successfully. An informal user experiment was finally conducted with four participants (aged 14–25), who tested the app in the field with both Android and iOS. One experimenter told the tasks to be carried out and a think-aloud protocol was adopted to gather users' comments. The application resulted easy to learn and to use, and only some minor problems emerged, which were easily fixed. All users appreciated the interaction experience with the app.

5 Conclusion

This paper describes the first version of a system supporting the creation of museum apps based on augmented reality and low-cost wireless devices. As to future work, we are planning to develop our own image recognition module, in order to avoid the use of Wikitude services and thus further reducing the usage costs. We are also thinking to include a speech recognition module to make interaction easier, especially for visually impaired people. Other accessibility barriers will also be taken into account in the future, such as cognitive disabilities.

An extended experimentation with several participants and different stakeholders, including people with different disabilities, is obviously necessary to demonstrate the acceptability and feasibility of the system (both app and CMS), and the financial viability of the idea.

We are also planning to extend the personalization feature by considering the possibility for museum curators to create different kinds of profiles according to the specific context, since in the current version the three profiles (child, tourist and scholar) are built-in in the CMS. This will require re-designing the CMS according to meta-design principles [6, 9] and make available end-user development facilities [14, 15] to curators, as discussed in [1, 5, 10]. In particular, as to children (or students in general), different interacting possibilities could be supported on the basis of age, including gamification techniques and serious games.

Acknowledgments. The research herein presented has been carried out within the Interdepartmental Laboratory Brixia Accessibility Lab (BrAL) of the University of Brescia where researchers share their knowledge in different fields. This research has been funded by the project "Work, Wealth, Production, Productivity", Programme "Health & Wealth 2015" of the University of Brescia.

References

1. Ardito, C., Buono, P., Costabile, M.F., Lanzilotti, R., Piccinno, A.: End users as co-designers of their own tools and products. J. Vis. Lang. Comput. **23**(2), 78–90 (2012)
2. Arenghi, A., Agostiano, M.: Cultural heritage and disability: can ICT be the 'missing piece' to face cultural heritage accessibility problems? In: Gaggi, O., Manzoni, P., Palazzi, C., Bujari, A., Marquez-Barja, J.M. (eds.) GOODTECHS 2016. LNICSSITE, vol. 195, pp. 70–77. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61949-1_8
3. Barbieri, G., Celentano, A., Orsini, R., Pittarello, F.: Understanding art exhibitions: from audioguides to multimedia companions. In: Proceedings of International Conference on Distributed Multimedia Systems, San Francisco Bay, USA, pp. 250–255. Knowledge Systems Institute, Skokie, IL (2009)
4. Bimber, O., Raskar, R.: Spatial Augmented Reality: Merging Real and Virtual Worlds. A.K. Peters Ltd., Natick (2005)
5. Celentano, A., Maurizio, M.: An end-user oriented building pattern for interactive art guides. In: Costabile, M.F., Dittrich, Y., Fischer, G., Piccinno, A. (eds.) IS-EUD 2011. LNCS, vol. 6654, pp. 187–202. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21530-8_15
6. Costabile, M.F., Fogli, D., Mussio, P., Piccinno, A.: A meta-design approach to end-user development. In: Proceedings of 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2005), pp. 308–310. IEEE Press, New York (2005)
7. Council of Europe: Framework Convention on Value of Cultural Heritage for Society. Council of Europe Treaty Series 199 (2005). https://rm.coe.int/1680083746
8. Dieck, C.T., Jung, T.H.: Value of augmented reality of cultural heritage sites: a stakeholder approach. J. Destin. Mark. Manag. **6**(2), 110–117 (2017)
9. Fischer, G., Fogli, D., Piccinno, A.: Revisiting and broadening the meta-design framework for end-user development. In: Paternò, F., Wulf, V. (eds.) New Perspectives in End-User Development. LNCS, pp. 61–97. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60291-2_4

10. Fogli, D., Piccinno, A.: Co-evolution of end-user developers and systems in multi-tiered proxy design problems. In: Dittrich, Y., Burnett, M., Mørch, A., Redmiles, D. (eds.) IS-EUD 2013. LNCS, vol. 7897, pp. 153–168. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38706-7_12
11. Gervautz, M., Schmalstieg, D.: Anywhere interfaces using handheld augmented reality. Computer **45**(7), 26–31 (2012)
12. Hansen, T.R., Eriksson, E., Lykke-Olesen, A.: Mixed interaction space - designing for camera-based interaction with mobile devices. In: Proceedings of Human Factors in Computing Systems, Extended Abstracts (CHI 2005), pp. 1933–1936. ACM Press, New York (2005)
13. Khan, M., de Byl, P.: Technology intervention for the preservation of intangible cultural heritage ICH. Int. J. Knowl. Soc. Res. **4**(2), 1–8 (2013)
14. Lieberman, H., Paternò, F., Wulf, V. (eds.): End User Development. Springer, Dordrecht (2006). https://doi.org/10.1007/1-4020-5386-X
15. Paternò, F., Wulf, V. (eds.): New Perspectives in End-User Development. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60291-2
16. Pedersen, I.: Radiating centers: augmented reality and human-centric design. In: Proceedings of IEEE International Symposium on Mixed and Augmented Reality, Orlando, FL, USA, pp. 11–16. IEEE press, New York (2009)
17. Pedersen, I., Gale, N., Mriza-Babaei, P., Reid, S.: More than meets the eye: the benefits of augmented reality and holographic displays for digital cultural heritage. ACM J. Comput. Cult. Herit. **10**(2), 11:1–11:15 (2017)
18. Ridel, B., Reuter, P., Laviole, J., Mellado, N., Couture, N., Granier, X.: The revealing flashlight: interactive spatial augmented reality for detail exploration of cultural heritage artifacts. J. Comput. Cult. Herit. **7**(2), 6:1–6:18 (2014)
19. Schieder, T.K., Adukaite, A., Cantoni, L.: Mobile apps devoted to UNESCO world heritage sites: a map. In: Xiang, Z., Tussyadiah, I. (eds.) Information and Communication Technologies in Tourism 2014, pp. 17–29. Springer, Cham (2013). https://doi.org/10.1007/978-3-319-03973-2_2
20. Tomiuc, A.: Navigating culture. Enhancing visitor museum experience through mobile technologies - from smartphone to google glass. J. Media Res. **7**(3), 33–46 (2014)
21. Vlahakis, V., Ioannidis, N., Karigiannis, J., Tsotros, M., Gounaris, M., Stricker, D., et al.: Archeoguide: an augmented reality guide for archeological sites. IEEE Comput. Graph. Appl. **22**, 52–60 (2011)
22. Zöllner, M., Stricker, D., Bleser, G., Pastarmov, Y.: iTACITUS - novel interaction and tracking paradigms for mobile AR. In: Proceedings of International Symposium on Virtual Reality, Archaeology and Cultural Heritage, Brighton, UK, pp. 110–117 (2007)

Supporting Young High-Functioning ASD Individuals in Learning the Concept of Money

Serena Caria[1], Fabio Paternò[1], Carmen Santoro[1(✉)], and Valentina Semucci[2]

[1] HIIS Laboratory, CNR-ISTI, Via G. Moruzzi, 1, 56124 Pisa, Italy
{serena.caria,fabio.paterno,carmen.santoro}@isti.cnr.it
[2] Azienda USL Toscana Nord Ovest, Via G. Garibaldi, 198, 56124 Pisa, Italy
valentina.semucci@uslnordovest.toscana.it

Abstract. We describe the design of a game-based Web application aimed to support high-functioning individuals affected by Autism Spectrum Disorder in gaining skills that can help them to understand the concept of money and apply it in practical situations of life. In order to evaluate the effectiveness and usability of the games, a user study involving six medium/high-functioning ASD individuals in their teens and above was carried out. Preliminary results were encouraging and show the potential advantages of such a system for training end users on practical life skills.

Keywords: Accessibility · Autism Spectrum Disorder · Serious games Web

1 Introduction

Autism Spectrum Disorder (ASD) refers to a broad range of neurodevelopmental disorders characterized by difficulties with social communication and interaction as well as restricted, repetitive and stereotyped patterns of behaviour [1]. While the exact ASD causes are unknown, it is believed that both genetic and environmental factors play a role in its development [2]. The term "spectrum" refers to the wide range of symptoms and levels of disability in functioning that affected people could display. Across it, three levels of functioning (low, medium, and high) are identified according to the severity of the disorder and thus the extent to which quality of life is negatively impacted. People who are affected most are called "low functioning", and they have quite severe impairments in all the three areas of reciprocal social interaction, communication, and repetitive behaviour. On the other extreme there are those whose quality of life is impacted less ("High Functioning" or HF): although they have a close to normal IQ (some even exhibit exceptional skills in specific areas), and language development can be normal, they have reduced social relationships connected with difficulties in starting or maintaining a conversation, deficits in emotional expression and recognition, limited range of interests, as well as troubles with organizational skills and abstract thinking.

There are many evidences [3, 4] that interactive technologies can be valuable tools in supporting computer-based learning of the core problematic areas of ASD (e.g.

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 374–384, 2018.
https://doi.org/10.1007/978-3-319-76111-4_37

communication, affective and interaction skills). Indeed, currently there are several technological solutions for supporting autistic people [3]: it has been noted that many individuals on the spectrum have a natural affinity with computers [5] due to the predictable and repeatable nature of technology that can create controlled environments, and which thus appeals to those who feel relieved by stability and routine. In addition, people with autism have strong visual processing skills, making them good candidates for approaches such as Augmentative and Alternative Communication (AAC) [6] and video modeling [7]. However, in spite of the growing attention paid to developing assistive applications for autistic people, we noted that so far most tools mainly address children developmental disabilities (social, cognitive, emotional, motor) within the ASD spectrum (see e.g. [8]). This may be due to the fact that earliest interventions (even starting from childhood: a first diagnosis can usually be made by the age of two) can give the highest chances to improve the core behavioural symptoms of autism. Thus, most work has concentrated to this age range. Less attention has been dedicated so far to other groups of autistic individuals which might need to gain different types of skills. This is the case of ASD adolescents and early adults, especially those characterised by a high functioning level. For such people, interventions might be needed helping them to achieve a more autonomous management of practical problems of daily living, for instance manage money and purchasing things, which can involve non-trivial aspects in cognition (recognize currency notes), decision-making (decide whether the object and the cost are congruous) and even mathematical competences (calculate exact change). The motivation of the lower interest that this category of users has attracted so far may be the fact that their high functioning nature makes their condition less visible to society, often leaving them with no support for coping with their real life problems [9] during their transition to early adulthood. Such challenges, often connected with e.g. increased demands of social relationships, self-determination/self-efficacy and more independent living, make adolescence and young adulthood one of the most difficult developmental periods in the life of these individuals. This situation is also exacerbated by the fact that public services tend to decline for ASD individuals after they leave high school, which is in turn frequently associated with substantial reduction (or even absence) of daytime activities such as higher education or work, and a disappointingly reduced abatement of ASD symptoms for HFA (High Functioning Autism) individuals during those developmental periods [10].

To alleviate these issues, we have designed and developed with the support of relevant stakeholders and users a set of games aimed at helping HFA autistic adolescents and early adults to more autonomously manage their life, more specifically to learn money management and also associated mathematical skills, which are both much needed by these users [11, 12]. In particular, the goal is to make them understand the concept of money and how to apply it in practical situations so that they can carry out everyday activities related to it. The games have been organised in a responsive Web application with multiple difficulty levels to support gradual learning. In recent years there has been a sharp rise in the number of research work specifically developed for ASD population. However, as highlighted in [13], everyone diagnosed with the ASD disorder is remarkably different. Thus, on the one hand it is extremely difficult to make generalizations, on the other hand developing new software and technologies for this incredibly diverse population is really a challenge. In the autism spectrum scale of functioning levels, HFA people represent the sub-group who get the least severe form of disability, generally characterized by the absence of language and cognitive impairments and, as such, their disability can be less visible to others. We noted that many contributions have focused on children and their learning and developmental related issues [14] because earliest interventions have the highest possibility to achieve the most benefits. Less studies support high-functioning teenagers/early adults in better (and more autonomously) managing their everyday life by means of training them in skills that can be generalizable and transferable to practical everyday situations (e.g. shopping). A review of technological interventions benefiting ASD adolescents is reported in [15].

2 The Games

In order to gather relevant requirements, we had five meetings with a speech therapist involved in educational and therapeutic activities for autistic people in a local health centre. Thanks to her knowledge of needs of autistic people we were able to collect several requirements that have driven the co-design and development of the application. We started by identifying current gaps, focusing on topics and skills that are currently difficult to teach (or are not taught at all) to HFA people in their teens or above with traditional methods and which would greatly benefit from personalised and motivating computer-based exercises that these users can practice autonomously (e.g. without the support of the caregiver) even in their familiar settings. The approach used was iterative and we progressed from ideas just sketched out at the beginning, to more refined design concepts which were then implemented in prototypes that were discussed during such meetings. As such, the application was designed in a participatory manner: caregivers participated in all the phases of the application development (while end users were involved only during evaluation) and we progressively discussed and captured teachers' reactions to the application we were going to develop. In particular, during such meetings the speech therapist provided us with useful insights about typical challenges HFA users have to cope with when they confront with a real life activity such as managing money. She also pointed out typical difficulties such individuals encounter when interacting with a computer (e.g. difficulties with text comprehension, risk of distractions), as well consolidated strategies therapists use to avoid upsetting their cared users (e.g. avoid explicitly disagreeing with them or saying 'no' to them), with the goal of smoothly progressing towards the expected learning objective. In addition, we also discussed with the speech therapist the typical exercises that trainers usually provide to HFA users in order to improve skills they are in need of, and which currently are often carried out by interacting with paper-based tools with the help of caregivers.

Thus, the resulting application was organised in such a way to reflect the three main learning objectives that we identified (see Fig. 1, left part): recognise the main currency (banknotes and coins) denominations; learn how to manage money change; learn what can realistically be purchased with a specific amount of money. For simplicity, we only considered showing the front view of banknotes and coins.

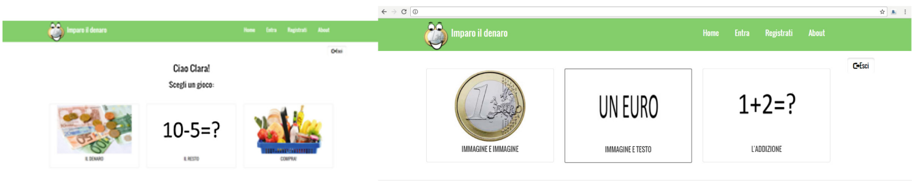


Fig. 1. Left: the three main games. Right: the three types of exercises included in the "Money" game

For consistency reasons some elements are common to all the exercises of the application. For instance, before the user starts to play the game, the application shows a text (accompanied by an audio) explaining what the user is supposed to do to solve that exercise. In addition, for all the games there is a "Help" button through which it is possible to activate a video that shows what to do for solving the game. In addition, in each page there are buttons to navigate through the pages of the application (e.g. go forward/backward), and a button to exit.

The "Money" game was developed to train people in recognizing different currency notes. When the user accesses this game, he can select one of three types of included games: "Image and image", "Image and text", and "Sum" (see Fig. 1, right part).

Image and Image. The first game (Image and Image) supports learning how to identify the main denominations and forms of the currency in terms of coins and banknotes. The technique used to support learning of money denominations is the *association*: the page is divided in two rows and the user has to drag each coin or banknote visualised in the bottom row on the corresponding similar coin or banknote shown in the top row (Fig. 2-left). If the user successfully completes the task, the associated coin/banknote is removed from the bottom row of the game and then the application shows a smiling yellow emoticon at the end of the exercise. If the user makes an error, a sad red face is shown. This game presents twelve levels of increasing difficulty i.e. the games become more challenging as soon as the user progresses in solving them. The level of difficulty is connected both to the number of elements to associate (i.e. the higher the number, the more difficult is the task) as well as to the similarity between the elements contained in the two rows. Once the user selects the

Fig. 2. Money game - *Image and Image* exercise (left); *Image and Text* exercise (right)

"Image and image" game, the user has to solve the first level (the simplest one), which presents only three images (one in the bottom row and two in the top row). Each level of this game presents a series of 4 games having the same difficulty. The second level presents overall five images (two in the bottom row and three in the top row). The last level includes six images (three at the bottom and three at the top).

Smiling faces have been included in all the games of the application, also accompanied by special sounds indicating either error or success. A yellow face appears only at the end of a successfully solved game, while a red face appears after each error. After the user makes two errors the position of the images on the bottom of the page is automatically shuffled so that the user cannot give answer only by memorizing the positions of the images without actually knowing the correct answer (this 'shuffle' strategy also applies to the next type of game).

Image and Text. While the previous game aimed to make people learn how to distinguish the various coins and banknotes, this game is aimed to make HFA students learn their names by using image-text associations: the user has to drag the coin or the banknote on its corresponding name (e.g. "10 euro"). As you can see from Fig. 2 (right), in this game decimals of euro were presented as "cents", while the currency unit was presented as "euro". This game has eight levels of complexity.

Sum. It aims to make users learn that, by combining coins and banknotes, it is possible to obtain further money values. Also in this case we used associations. In the top row the application shows a sequence of coins and banknotes with an in-between "+" sign, while at the end of this sequence, after the "=" symbol (see Fig. 3-left) the image of an open wallet is visualized, with two buttons available: one (initially de-activated) to try again the game, the other to signal that the exercise is finished.

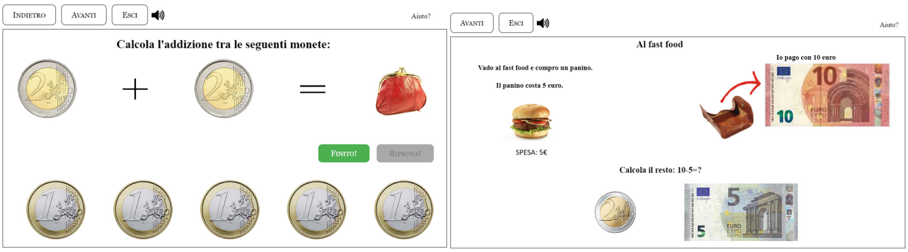


Fig. 3. Money game: sum exercise (left); money change (right)

To solve the game, the user has to drag in the wallet the precise amount of money needed to correctly solve the sum and then click on the button used to signal that the exercise is finished. If the exercise is correctly solved a smiling yellow face will appear, otherwise a red face will be shown and the "Retry" button will be enabled. Also in this case, various levels of difficulty have been designed, ranging from the simplest one (where only two coins are combined), to the most complex ones in which also banknotes are included.

The "Money Change" game is dedicated to learning the concept of money change. Differently from the previous games, at the beginning the application asks the user to select a video modeling [7] item of interest and, only after this, he can access the game. Video modeling items are videos depicting exemplary behaviour: the user is expected to observe a videotape of a model doing the task (or the skill) the teacher wishes to teach, and then such model is subsequently practiced and imitated. Video modeling is a behavioural technique which has been shown as being particularly valuable for autistic people. In our case two video modeling elements were prepared (one having a male actor and another one with a female one, to better suit the preference of the user), to show how potential customers should behave to buy items in a stationery store. During such videos some 'focus' elements were also used to emphasise key steps/objects on which users should focus their attention ('focus' objects are generally rendered in the video at a reduced speed and by zooming in them). In the videos we built, the object of a specific focus was a pen that a customer wants to buy. After having watched the video, the user could access the game, which was presented in a problem-like manner. The application shows some scenarios of everyday life, and in each of them there is a purchasing action implying a change: the user has to solve the mathematical subtraction associated with the considered scenario (see Fig. 3-right) by means of selecting, among the images available in the bottom part of the user interface, the one representing the right solution. Also this game presents eight levels of difficulty.

The "Buy It!" game allows the student to learn what can be realistically purchased with a specific amount of money, by simulating a real situation. At the start of the game the user is presented with a "What can I buy with…" string in the top part of the UI, as well as the amount of money considered (the question is further emphasized by an emoticon showing a thinking face, see Fig. 4).

Fig. 4. The *Buy it!* game

Then the user has to select among three different images (shown in the bottom part of the UI) the one that should correspond to the indicated amount of money. This game is the last one in the series of exercises since it implies that the user has already got the skills needed to solve the previous ones. It provided different levels of difficulty, smiling faces to indicate success/error, and a shuffling strategy applied after two user's errors.

3 Evaluation

We tested the application to assess to what extent it helps users in understanding the concepts associated with money management. The test was articulated in three steps: we first submitted a pre-test questionnaire to gather demographic data; then the participants had to carry out five tasks (each task associated with a type of game) using the application; finally, they had to fill in a System Usability Scale (SUS) questionnaire [16], for measuring the usability of the application, which consists of a 10 item questionnaire with five response options for respondents, from “strongly agree” to “strongly disagree”.

We tested the application at the premises of an association supporting ASD individuals. The test involved six male ASD individuals. In the following, we use a set of pseudonyms to indicate them: Francesco, Jacopo, Gabriele, Giulio, Andrea, Mattia). All of them had been diagnosed with medium/high functioning ASD. Their age ranged from 16 to 22 years (M = 18.5; SD = 2.2). In order to better manage the participants during the evaluation, we divided them into two groups consisting of 3 members each, according to their age. One group (Francesco, Jacopo and Gabriele) is composed of three early adults (age range 20–22), all having a high school diploma. Their level of use of technological devices is pretty high, the most used device is the smartphone, exploited for browsing the Web. As for tablets and PCs, two users use both devices, while one user never used tablets and he uses the PC just few times a week. All of them use such devices to browse the Web and play games.

The three remaining users (Giulio, Andrea and Mattia) are teenagers (16–17 years old) still finishing secondary school. Their familiarity with devices was rather varied. While one user was particularly familiar with technology (having even some knowledge of JavaScript), another user occasionally uses the smartphone (e.g. to call mother) and he did not have any experience with PCs. The last user had low familiarity with smartphones (used only for photos and Web browsing) and tablets, and very good familiarity with PCs (used for Web browsing and games).

The tasks assigned to users were to access and complete every level of the five games developed in the application (Task1: Money/Image and Image game; Task2: Money/Image and Text game; Task3: Money/Sum game; Task4: Change game; Task5: Buy it! game). Participants had to use twice the application: the second try was carried out one week after the first one. This was done to understand whether any improvements in using the application could be detected over time. For evaluation goals, the web application was enhanced with a logging tool implemented in JavaScript. The test was done by using a Windows-based laptop having a 15.6” monitor with a 1366 $\times$ 768 resolution, and a AMD Quad-Core A8-6410 processor. Each participant was allowed to interact with each exercise for at maximum 5 min. Since they had to solve five tasks, the total duration of the study for each participant was about half an hour.

Task Success. The “Task success” metric is used to verify whether and how users completed the assigned tasks.

Four success levels were identified: we assigned a “1” score if the user did not have any problem, “2” when just a small problem was found (e.g. two slight errors occurred and then the user was able to continue the interaction), “3” when the user made more severe errors, and "4" when the user was not able at all to solve the exercise. Figure 5 shows two stacked bar charts indicating the success level for each task in both tests: during the second try we had higher scores than in the first one. In particular, in the second trial, we had a higher number of tasks successfully carried out by users, and the number of errors due to major problems were less than in the first trial.

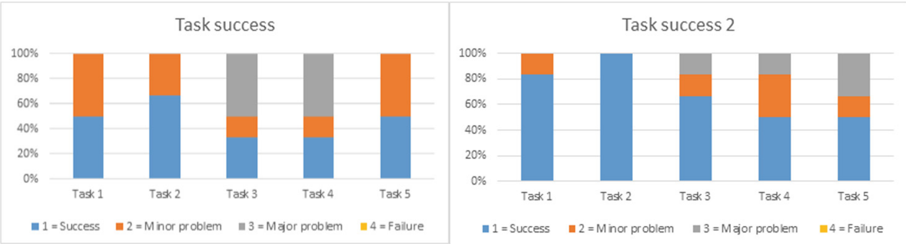


Fig. 5. Task success – first test (left) and second test (right)

Time on Task. Data related to this metric was collected by using a logging script which was used to enhance the application exploited during the evaluation. We gave a maximum of five-minute time to solve each task and then we verified whether in this interval of time the user was able to solve the exercise. If not, the evaluator recorded the last level reached by the user. Levels reached during the two trails were then compared each other. In this comparison, other factors were also considered e.g. errors done and requests for help. The results on this metric gathered in both trails are shown in Fig. 6 (left and right), where the X axis refers to tasks and Y axis refers to time on task in seconds. Lower and upper bounds were also calculated on such data: apart from one case associated with lower bound values on Task 2 (namely: 72 s in the second session and 61.2 s in the first session), time on task overall improved among the two sessions.

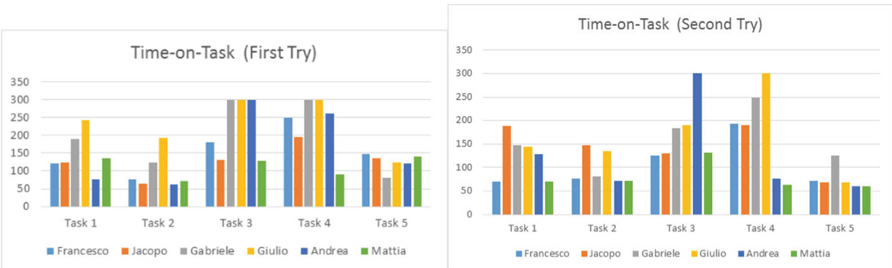


Fig. 6. *Time on Task* in the first test (left) and in the second test (right)

Errors. We also counted the number of errors done by participants and also in this case, overall, we had improvements. Apart from one user (Francesco, who in the first trial made just two errors in the whole test, while in the second session made two errors in the last three tasks), such number decreased between the two evaluation sessions (see Fig. 7 left and right).

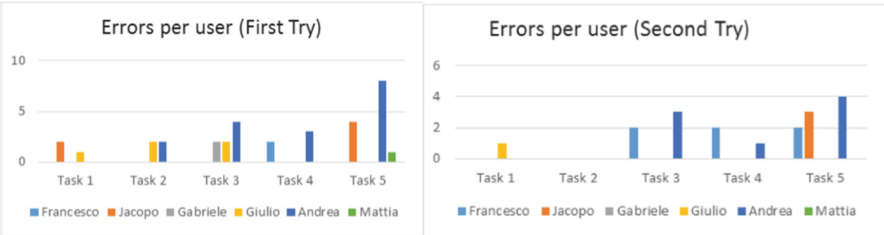


Fig. 7. Errors in the first test (left) and in the second test (right)

Satisfaction. In order to assess user satisfaction, we considered the SUS questionnaire, which was filled in by users at the end of each session. According to Jacopo, Francesco and Andrea, the application was not very usable (they got 50, 52.5 and 60 scores respectively), whereas the other three users judged it usable (the scores were 100, 85, 77.2). To this regard, the moderator noticed that, while filling in the questionnaire, users seemed not much motivated and also a bit tired. In general, users appreciated the application. In the first trial the group composed of Francesco, Jacopo and Gabriele highlighted difficulties with calculations (some of them used the calculator during the test). Another problem was connected with reading the textual strings within the "*Image and Text*" game, which particularly affected the performance of one user. One user indicated the need of having longer time to solve the games than the allowed one (5 min). Another user complained about the smiling faces used in the games. Users found the "*Image and Image*" and "*Image and Text*" games as too simple, whereas the "*Sum*" game was found a bit too difficult. During the second session some of them showed some progress.

Regarding the other group (Giulio, Andrea and Mattia), contrasting judgements came from them in the first trial. Two of them evaluated the application nice yet boring, whereas one found the application both valuable and stimulating. Giulio and Andrea had problems with calculations (in the *Sum* and *Money Change* games). Mattia had just one concern in the *Buy it!* game as he thought that with the same amount of money it is possible to buy different things in the game. Mattia did also not appreciate much the provided audio feedback. During the second test, two of them showed evident performance improvements.

4 Conclusions and Future Work

In this paper we present a game-based Web application aimed at supporting high functioning ASD people in their teens and above to gain practical life skills connected with money management. A user study was conducted with six high functioning ASD individuals to assess its effectiveness and usability. Overall, the test shows encouraging results in the potentiality of training high functioning ASD individuals in acquiring skills regarding money management. In addition, being a Web application, it can be autonomously exploited by users whenever they want, and at their own pace. The results are promising although additional work to make the application more personalised and adaptive to user's needs, preferences and behaviour should be done in the future, as well as further empirical studies to better evaluate it.

References

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders (5th ed.) (2013)
2. National Institute of Neurological Disorders and Stroke. Autism Spectrum Disorder Fact Sheet. https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Fact-Sheets/Autism-Spectrum-Disorder-Fact-Sheet
3. Aresti-Bartolome, N., Garcia-Zapirain, B.: Technologies as support tools for persons with autistic spectrum disorder: a systematic review. Int. J. Environ. Res. Public Health **11**(8), 7767–7802 (2014). https://doi.org/10.3390/ijerph110807767
4. Bossavit, B., Parsons, S.: Designing an educational game for and with teenagers with high functioning autism. In: Bossen, C., Smith, R.C., Kanstrup, A.M., McDonnell, J., Teli, M., Bødker, K., (eds.) Proceedings of the 14th Participatory Design Conference: Full papers, (PDC 2016), vol. 1, pp. 11–20. ACM, New York (2016). https://doi.org/10.1145/2940299.2940313
5. Moore, D.: Computers and people with autism/asperger syndrome, in Communication (the magazine of the National Autistic Society), pp. 20–21 (1998)
6. Cafiero, J.: Meaningful Exchanges for People with Autism: An Introduction to Augmentative & Alternative Communication. Woodbine House, Bethesda (2005)
7. Delano, M.E.: Video modeling interventions for individuals with autism. Remedial Spec. Educ. **28**(1), 33–42 (2016)
8. Bernardini, S., Porayska-Pomsta, K., Smith, T.J.: ECHOES: an intelligent serious game for fostering social communication in children with autism. J. Inf. Sci. **264**, 41–60 (2014)
9. Hong, A.P.C.I., van Heugten, S., Kooken, T., Vinke, N., Vromans, M., Shahid, S.: The augenda: structuring the lives of autistic teenagers. In: ASSETS 2010, pp. 283–284. ACM (2010). https://doi.org/10.1145/1878803.1878872
10. Shattuck, P.T., Wagner, M., Narendorf, S., Sterzing, P., Hensley, M.: Post-high school service use among young adults with an autism spectrum. Arch. Pediatr. Adolesc. Med. **165** (2), 141–146 (2011)
11. Waters, H.E., Boon, R.T.: Teaching money computation skills to high school students with mild intellectual disabilities via the touchmath© program: a multi-sensory approach. Educ. Train. Autism Dev. Disabil. **46**(4), 544–555 (2011)
12. Whitby, P.J.S.: Teaching mathematics to students with high functioning autism. In: A Guide to Teaching Students with Autism Spectrum Disorders (2013)
13. Morris, R.R., Kirschbaum, C.R., Picard, R.W.: Broadening accessibility through special interests: a new approach for software customization. In: Proceedings of ASSETS 2010, pp. 171–178. ACM (2010). https://doi.org/10.1145/1878803.1878834
14. Ploog, B.O., Scharf, A., Nelson, D., et al.: Use of computer-assisted technologies (CAT) to enhance social, communicative, and language development in children with autism spectrum disorders. J. Autism Dev. Disord. **43**, 301 (2013). https://doi.org/10.1007/s10803-012-1571-3

15. Odom, S.L., Thompson, J.L., Hedges, S., et al.: Technology-aided interventions and instruction for adolescents with autism spectrum disorder. J. Autism Dev. Disord. **45**, 3805 (2015). https://doi.org/10.1007/s10803-014-2320-6
16. Brooke, J.: SUS: a “quick and dirty” usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, A.L. (eds.) Usability Evaluation in Industry. Taylor and Francis, London (1996)

PdUC-D: A Discretized UAV Guidance System for Air Pollution Monitoring Tasks

Oscar Alvear[1,2(✉)], Carlos T. Calafate[1], Nicola Roberto Zema[3], Enrico Natalizio[4], Enrique Hernández-Orallo[1], Juan-Carlos Cano[1], and Pietro Manzoni[1]

[1] Department of Computer Engineering, Universitat Politècnica de València, Camino de Vera, S/N, 46022 Valencia, Spain
oscar.alvear@alttics.com

[2] Department of Electrical Engineering, Electronics and Telecommunications, Universidad de Cuenca, Av. 12 de Abril, S/N, Cuenca, Ecuador

[3] Univ Lille Nord de France, IFSTTAR, COSYS, 59650 Villeneuve d'Ascq, France

[4] Sorbonne Universités, Université de Technologie de Compiègne, CNRS, Laboratoire Heudiasyc, 57 Avenue de Landshut, CS 60319, 60203 Compiegne Cedex, France

Abstract. Discretization is one of the most efficient mathematical approaches to simplify (optimize) a system by transforming a continuous domain into its discrete counterpart. In this paper, by adopting space discretization, we have modified the previously proposed solution called PdUC (Pollution-driven UAV Control), which is a protocol designed to guide UAVs that monitor air quality in a specific area by focusing on the most polluted areas. The improvement proposed in this paper, called PdUC-D, consists of an optimization whereby UAVs only move between the central tile positions of a discretized space, avoiding to monitor locations separated by small distances, and whose actual differences in terms of air quality are barely noticeable. Experimental results show that PdUC-D drastically reduces convergence time compared to the original PdUC proposal without loss of accuracy.

Keywords: UAV control system · Air pollution monitoring
Discretized system

1 Introduction

Air pollution is a hazard not only affecting urban areas (cities) [1], but also rural and industrial environments [2] in different aspects such as crop yield, forest and animal health, among others.

In the literature, we can observe that traditional methods for air pollution monitoring (fixed monitoring stations) are gradually being replaced by mobile crowdsensing sensors that are small enough to be carried around by users, or installed in different vehicles like taxis, buses, bicycles, or any type of land vehicle [3–7].

B. Guidi et al. (Eds.): GOODTECHS 2017, LNICST 233, pp. 385–394, 2018.
https://doi.org/10.1007/978-3-319-76111-4_38

The crowdsensing approach is not feasible in rural areas because it clearly requires a minimum number of sensors to be moving inside the target area to be applicable, a requirement that is typically not met in these remote environments. For instance, in this type of scenarios, vehicular traffic is quite scarce, being limited to the main transportation arteries, thereby failing to provide the required granularity in both time and spatial domains.

To effectively carry out monitoring tasks in rural scenarios, an attractive option is to use Unmanned Aerial Vehicles (UAVs) equipped with commercial off-the-shelf (COTS) sensors, allowing them to act as mobile sensors, able to reach poorly accessible areas [8]. In fact, this approach allows monitoring most locations in any target area due to UAV flexibility and maneuverability, like the capability to take samples while hovering.

Focusing on UAV control systems for air pollution monitoring tasks, we have previously noticed that there are no systems optimized for these purposes. So, we proposed PdUC (Pollution-driven UAV Control [9]), a solution that puts focus on the most polluted regions by combining a chemotaxis metaheuristic with adaptive spiral mobility patterns to automatically track pollution sources and surrounding pollution values in a given target area. In that previous work [9] we showed that PdUC achieves better performances than standard mobility approaches, like the Spiral and the Billiard patterns, in terms of discovering the most polluted areas in a shorter time span. In this paper we propose an optimized algorithm called PdUC-D, which is based on PdUC, but applies space discretization to substantially reduce the convergence time while achieving similar levels of accuracy.

This paper is organized as follows: in Sect. 2 we describe the proposed PdUC-D protocol. Section 3 presents the implementations of the protocol in the R tool, along with a performance comparison against the original PdUC protocol. Finally, in Sect. 4, we present the conclusions of our work and the future work.

2 PdUC-D: Discretized Pollution-Driven UAV Control Protocol

Despite PdUC [9] is more effective than other mobility patterns (Spiral and Billiard) in terms of polluted areas monitoring times, finding the most highly polluted locations earlier, PdUC still spends too much time focusing on small variations (variations produced by sensor errors, or little pollution variations) in nearby areas, which are not too useful when obtaining the global pollution map; on the contrary, the Spiral and Billiard models present simpler mobility patterns that, by themselves, avoid such redundant sampling. So, in this work, we attempt to avoid redundant movements (sampling) by discretizing the target area, dividing it into small tiles.

The main idea is to optimize PdUC by discretizing the whole target area in a grid forming small tiles, as shown in Fig. 1. The UAV can only move to the center of each tile, thereby reducing redundant sampling, which in turn reduces the full coverage time significantly. Each tile is monitored only once.

Fig. 1. Example of a discretized area, calculating the tiles and their center to restrict movements.

PdUC-D, just like PdUC, is based on a *chemotaxis* metaheuristic and on an adaptive spiral, with the difference that now both these mechanisms are adapted to operate with discretized space environments. Therefore, PdUC-D first starts by searching the tile with the highest pollution level (Search phase). Next, it covers the surrounding area following an adaptive spiral until all the area is covered, or until it can find another tile with a higher pollution value (Explore phase), thereby switching back to the Search phase.

We modify the PdUC phases by adapting its functionality to space discretization as follows:

The search phase is based on a chemotaxis mobility pattern, as shown in Fig. 2: a particle moving in an euclidean plane between two tiles, and following a specific direction, moves towards the next tile in the same direction (*Run* move) if the pollution variation is increasing along it. Otherwise, if the pollution variation is decreasing, it moves around the tile with higher previously monitored pollution values, assigning higher priority to nearer tiles (*Tumble* move); namely, it chooses the nearest tile. If all tiles around the one with the highest detected value have already been monitored, the algorithm switches to the *Explore* phase.

The *Explore* phase is based on a adaptive spiral, as shown in Fig. 3 (left): Starting at the tile with the highest monitored pollution value, it then follows a square spiral. For each round in the spiral, it skips an increasing number of tiles. Namely, in the first round it has an radius of 3 tiles and skips 1 tile; in the second round, it has an radius of 5 tiles and skips 2 tiles, and so on. To avoid

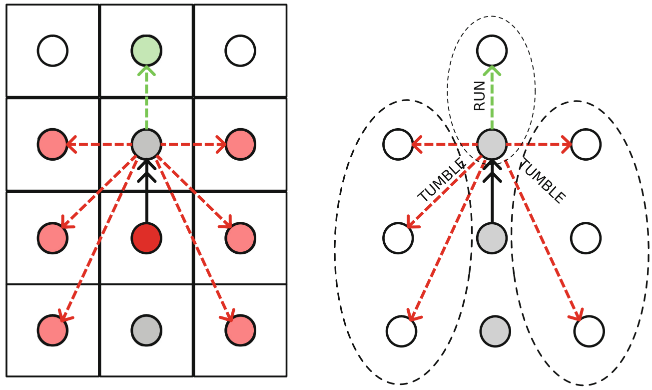


Fig. 2. Search phase: calculation of the next tile following the chemotaxis-based mobility pattern.

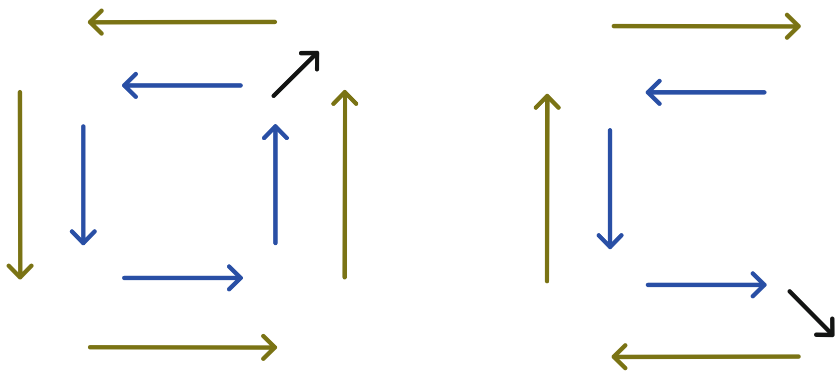


Fig. 3. Explore phase: adaptive spiral followed by PdUC-D showing a normal spiral movement (left) and alternating movement (right).

excessively long steps, if the spiral radius reaches a scenario border, or previously monitored areas, the direction of the spiral is altered, as shown Fig. 3 (right).

Figure 4 shows the behaviour of the adaptive spiral: it first starts by following a square spiral but, when it reaches the border, it alternates the direction of movement to rotate in the opposite direction.

With regard to movement control, and to avoid previously monitored areas, we use two matrixes: $P_{m,n}$, and $B_{m,n}$ to store the sampled values and the monitored tiles, respectively. Notice that $n \times m$ represent the size of the grid, and the position of the tiles.

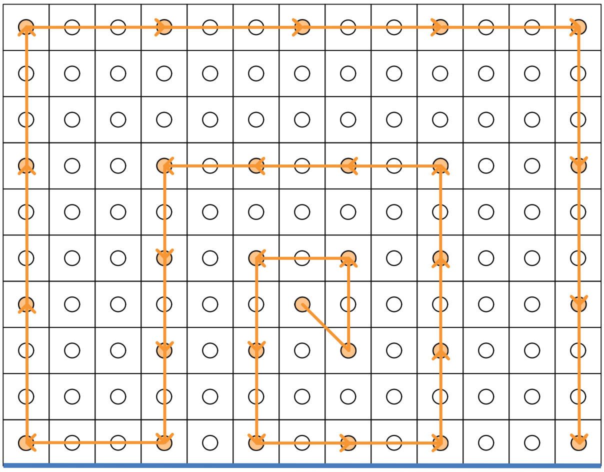


Fig. 4. Explore phase: example of the adaptive square spiral showing the possible moves.

$$P_{m,n} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,n} \end{pmatrix} B_{m,n} = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{pmatrix} \quad (1)$$

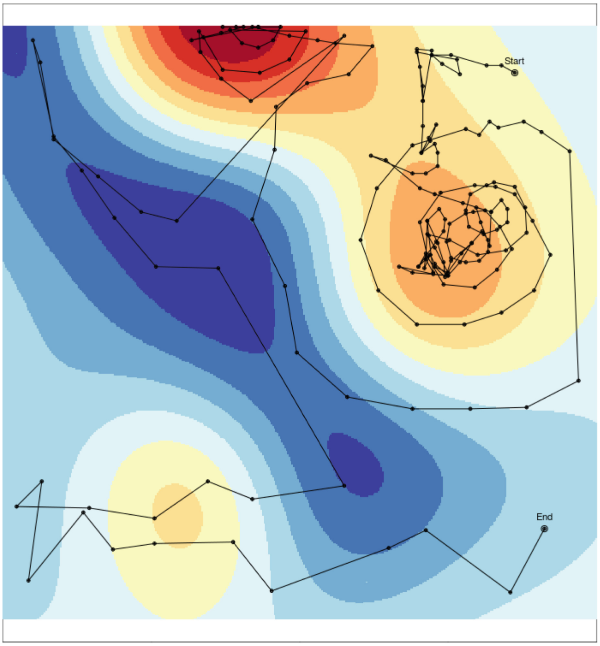
First, both matrices are initialized, P with NA (null), and B with 0's. In the *search* phase, when monitoring a tile $t_{x,y}$, the pollution values are stored in $P_{x,y}$, and $B_{x,y}$ are set to 1. In the *explore* phase, as well as in the *search* phase, when monitoring a tile, both P and B values are stored but, when finishing a spiral round, all tiles inside the square are set as visited in B, thereby avoiding to monitor the same area again in the future.

3 Validation

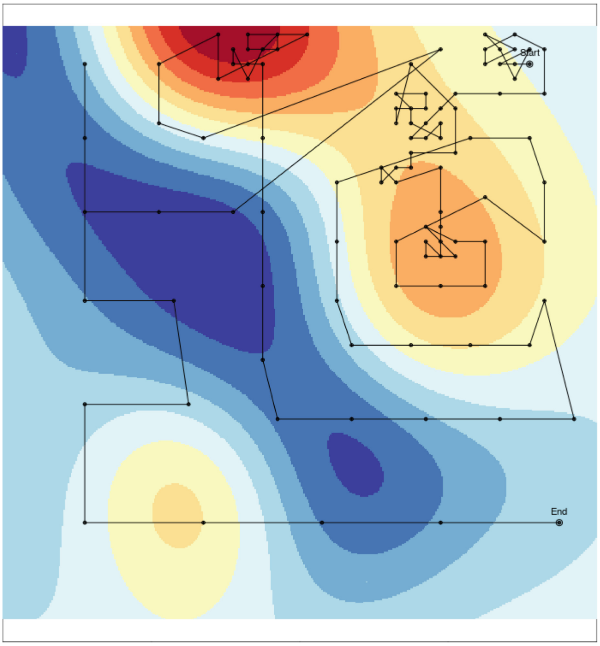
We have implemented PdUC-D in the *R Graph* tool [10], and we have run several simulations with different configurations.

To prepare a suitable data environment, we have created various pollution distribution maps representing ozone levels to be used as inputs for testing. These pollution maps were also generated using the *R Graph* tool following Kriging-based interpolation [11]. In particular, a Gaussian distribution is used to adjust the parameters coming from random data sources of ozone concentration. The actual values range between 40 and 180 ppb (parts-per-billion), thereby providing a realistic ozone distribution.

Obtained data using PdUC-D was compared against previous results obtained using PdUC [9]. Figure 5 shows an example of the path followed by an



(a) PdUC



(b) PdUC-D

Fig. 5. Example of an path followed by an UAV guided by PdUC and PdUC-D protocols.

UAV using (a) PdUC and (b) PdUC-D as a guidance system. As expected, both algorithms have, in general, a similar behaviour: the UAV starts a search process throughout the scenario until it locates a position with the highest degree of pollution (local maximum). Afterward, it follows a spiral pattern to gain awareness of the surrounding gradients. If, while following the spiral-shaped scan path, it finds a higher pollution value, the algorithm again switches to the search phase. Finally, when the entire target area has been sampled, the algorithm finishes. When adopting PdUC-D, though, we can clearly see that it achieves better performance while avoiding redundant sampling.

To compare PdUC-D against PdUC [12] we use the same simulation parameters, previously used for validating PdUC. Table 1 summarizes the parameters used in the simulations.

Table 1. Simulation parameters.

Parameter	Value
Area	4×4 Km
Pollution range	[40–180] ppb
Sampling error	10 ppb
Max. speed	20 m/s
Sampling time	4 s
Step distance	100 m
Mobility models	Billiard, Spiral and PdUC

Since we are proposing the PdUC and PdUC-D algorithms for rural environments, the simulation area defined is a 4×4 Km area. The pollution distribution relies on synthetic maps that are generated by combining a random Kriging interpolation following a Gaussian model with values between 40 and 180 units based on the Air Quality Index (AQI) [13]. Since samples are taken using off-the-shelf sensors, which are not precise, we introduce a random sampling error of ± 10 ppb based on real tests using the MQ131 (Ozone) sensor. In our simulation, we set the maximum UAV speed to 20 m/s, a value achievable by many commercial UAVs. The step distance defined between consecutive samples is 100 m since it offers a good trade-off between granularity and flight time. Once a new sampling location is reached, the monitoring time per sample is defined to be 4 s.

Figure 6 shows the Cumulative Distribution Function relative to the time required to cover the whole area for PdUC and PdUC-D mobility models. It can be seen that the PdUC-D model spends much less time (1500–3000 s) than the PdUC model (1800–4300 s) to achieve the same goal.

To gain further insight into the goodness of the proposed algorithm, we also analyze the relative error for the two mobility models at different time instants (600, 1200, 1800, 2400, 3000 and 600 s); this error is defined by Eq. 2:

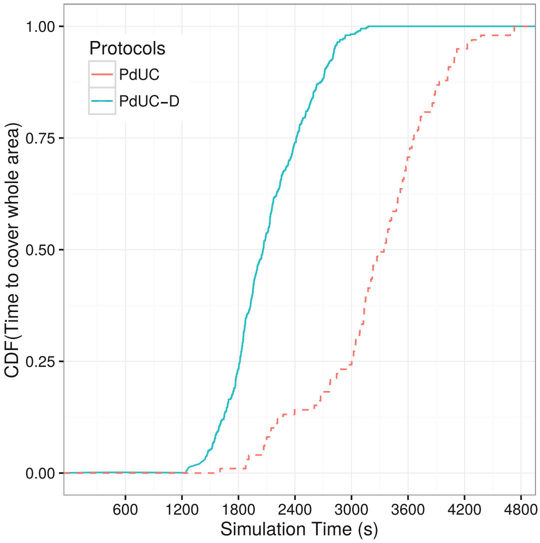


Fig. 6. Cumulative Distribution Function of the time spent at covering the complete area for the PdUC and PdUC-D mobility models.

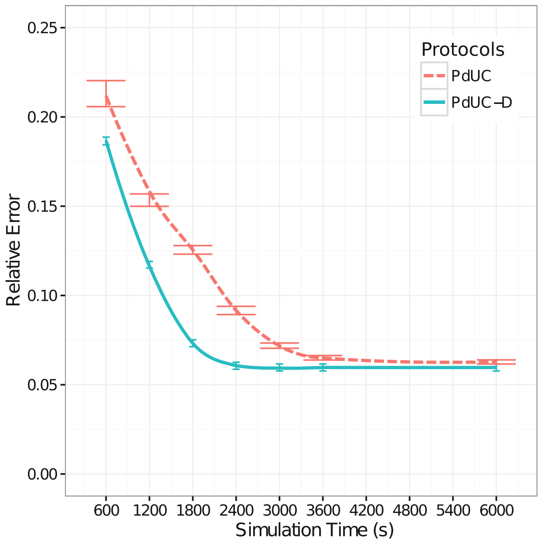


Fig. 7. Relative error comparison between the PdUC and PdUC-D mobility models at different times.

$$e_t = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} |\frac{s_{x,y,t} - b_{x,y}}{\triangle b}|}{m \cdot n} \tag{2}$$

where, e_t is the relative error at time t; $s_{x,y,t}$ is the recreated pollution value at position (x, y) using the samples taken during simulation until time t, $b_{x,y}$ is the reference pollution value at position (x, y), and n and m are the dimensions of the target area, respectively.

Figure 7 shows the temporal evolution of the relative error between model-based predictions (PdUC and PdUC-D) and the reference values. We can observe that both mobility models have roughly the same behavior: they start with a high relative error, which is foreseeable since we are using Kriging interpolation to recreate the pollution distribution, and it tends to the mean value when the number of samples is not enough. Then, as more samples become available, the spatial interpolation process quickly becomes more precise. Moreover, we can observe that, even in this analysis, PdUC-D obtains better results than PdUC by significantly reducing the relative error at different times.

4 Conclusions and Future Work

Air pollution monitoring in rural areas is a relevant issue that typically finds many obstacles due to the lack of monitoring infrastructures, and due to the complexity of having ground mobile sensors in many cases. In this context, UAVs equipped with air quality sensors emerge as a novel and powerful alternative.

In this paper we follow this assumption by describing an algorithm for air pollution monitoring tasks called PdUC-D (Discretized Pollution-driven UAV Control), that is based on a previous work (PdUC). In particular, it operates as an UAV guidance system to move towards the most polluted areas, and map pollution in the surrounding area. PdUC-D has the same phases as the original PdUC proposal (Search and Explore), and it is based in the same principles (Chemotaxis and Adaptive Spiral), but its functionality was modified to work in an space-discretized area, thereby making it much more optimal.

We have compared PdUC-D against PdUC by creating several simulations in the R Tool, and comparing these results with the previously obtained ones. Experimental results show that PdUC-D has much better performance than PdUC in all aspects, reducing the time to cover a same area, and reducing the error.

The next step of our research is to translated our algorithm to a real UAV, and test it in real-world testbed.

Acknowledgment. This work has been partially carried out in the framework of the DIVINA Challenge Team, which is funded under the Labex MS2T program. Labex MS2T is supported by the French Government, through the program "Investments for the future" managed by the National Agency for Research (Reference: ANR-11-IDEX-0004-02). This work was also supported by the "Programa Estatal de Investigación, Desarrollo e Innovación Orientada a Retos de la Sociedad, Proyecto I+D+I TEC2014-52690-R", the "Programa de becas SENESCYT de la República del Ecuador", and the Research Direction of the University of Cuenca.

References

1. Seaton, A., Godden, D., MacNee, W., Donaldson, K.: Particulate air pollution and acute health effects. Lancet **345**(8943), 176–178 (1995)
2. McFrederick, Q.S., Kathilankal, J.C., Fuentes, J.D.: Air pollution modifies floral scent trails. Atmos. Environ. **42**(10), 2336–2348 (2008)
3. Alvear, O., Zamora, W., Calafate, C., Cano, J.-C., Manzoni, P.: An architecture offering mobile pollution sensing with high spatial resolution. J. Sens. **2016** (2016)
4. Adam-poupart, A., Brand, A., Fournier, M., Jerrett, M., Smargiassi, A.: Spatiotemporal modeling of ozone levels in Quebec (Canada): a comparison of kriging, land-use regression (LUR), and combined Bayesian maximum entropy'LUR approaches. Environ. Health Perspect. **970**(January 2013), 1–19 (2014)
5. Pujadas, M., Plaza, J., Teres, J., Artıñano, B., Millan, M.: Passive remote sensing of nitrogen dioxide as a tool for tracking air pollution in urban areas: the madrid urban plume, a case of study. Atmos. Environ. **34**(19), 3041–3056 (2000)
6. Eisenman, S.B., Miluzzo, E., Lane, N.D., Peterson, R.A., Ahn, G.-S., Campbell, A.T.: Bikenet: a mobile sensing system for cyclist experience mapping. ACM Trans. Sens. Netw. (TOSN) **6**(1), 6 (2009)
7. André, M.: The artemis European driving cycles for measuring car pollutant emissions. Sci. Total Environ. **334**, 73–84 (2004)
8. Alvear, O., Calafate, C.T., Hernández-Orallo, E., Cano, J.-C., Manzoni, P.: Mobile pollution data sensing using UAVs. In: The 13th International Conference on Advances in Mobile Computing and Multimedia (2015)
9. Alvear, O.A., Zema, N.R., Natalizio, E., Calafate, C.T.: A chemotactic pollution-homing UAV guidance system. In: 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 2115–2120. IEEE (2017)
10. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2016)
11. Stein, M.L.: Statistical Interpolation of Spatial Data: Some Theory for Kriging. Springer, New York (1999). https://doi.org/10.1007/978-1-4612-1494-6
12. Alvear, O., Zema, N.R., Natalizio, E., Calafate, C.T.: Using UAV-based systems to monitor air pollution in areas with poor accessibility. J. Adv. Transp. **2017** (2017)
13. United States Environmental Protection Agency. Air Quality Index Available (2015) http://cfpub.epa.gov/airnow/index.cfm?action=aqibasics.aqi

Author Index