CENTRIC 2015

The Eighth International Conference on Advances in Human oriented and Personalized Mechanisms, Technologies, and Services

ISBN: 978-1-61208-440-4

November 15 - 20, 2015

Barcelona, Spain

CENTRIC 2015 Editors

Lasse Berntzen, Vestfold University College - Tønsberg, Norway

Stephan Böhm, RheinMain University of Applied Sciences - Wiesbaden, Germany

CENTRIC 2015

Forward

The Eighth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services (CENTRIC 2015), held on November 15 - 20, 2015 in Barcelona, Spain, addressed topics on human-oriented and personalized mechanisms, technologies, and services, commonly known as I-centric.

There is a cohort of technologies that favored the so called “user-centric” services and applications. While some of them reached some maturity, others are to prove their economics (WiMax, IPTV, RFID, etc). The human-oriented and personalized technologies and services rely on a key set of features, some to be deployed, others getting more mature (personal profiles, preferences, identity, proximity, personal devices, etc.). Following, advanced applications covering human related activities benefit from personalized and human-oriented networks and services, especially preventive and personalized medicine, body networks and devices, or anticipative systems.

The conference provided a forum where researchers were able to present recent research results and new research problems and directions related to them. The conference sought contributions presenting novel result and future research in all aspects of user-centric mechanisms, technologies, and services.

Similar to the previous editions, this event continued to be very competitive in its selection process and very well perceived by the international community. As such, it attracted excellent contributions and active participation from all over the world. We were very pleased to receive a large amount of top quality contributions.

We take here the opportunity to warmly thank all the members of the CENTRIC 2015 technical program committee as well as the numerous reviewers. The creation of such a broad and high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and efforts to contribute to the CENTRIC 2015. We truly believe that thanks to all these efforts, the final conference program consists of top quality contributions.

This event could also not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the CENTRIC 2015 organizing committee for their help in handling the logistics and for their work that is making this professional meeting a success.

We hope the CENTRIC 2015 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in personalization research. We also hope Barcelona provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

CENTRIC 2015 Chairs

CENTRIC 2015 General Chairs

Lasse Berntzen, Vestfold University College - Tønsberg, Norway

Stephan Böhm, RheinMain University of Applied Sciences - Wiesbaden, Germany

CENTRIC Advisory Chairs

Michael Gebhart, iterate GmbH, Germany
Vitaly Klyuev, University of Aizu, Japan
Mario Hoffmann, Fraunhofer-Institute for Secure Information Technology (SIT) - Darmstadt, Germany
Guadalupe Ortiz, University of Cádiz, Spain
Thomas Marx, University of Applied Sciences Bingen, Germany

CENTRIC 2015 Research Institute Liaison Chairs

R. Hariprakash, Arulmighu Meenakashi Amman College of Engineering, India
Muhammad Iftikhar, University of Malaysia Sabah, Malaysia
Kong Cheng, Telcordia Technologies - Piscataway, USA
Gottfried Zimmermann, Stuttgart Media University, Germnay

CENTRIC 2015 Industry/Research Chairs

Eve Maler, Forrester Research, USA
Rainer Falk, Siemens AG - München, Deutschland
Juong-Sik Lee, Nokia Research Center - Palo Alto, USA

CENTRIC 2015

Committee

CENTRIC 2015 General Chairs

Lasse Berntzen, Vestfold University College - Tønsberg, Norway
Stephan Böhm, RheinMain University of Applied Sciences - Wiesbaden, Germany

CENTRIC Advisory Chairs

Michael Gebhart, iteratec GmbH, Germany
Vitaly Klyuev, University of Aizu, Japan
Mario Hoffmann, Fraunhofer-Institute for Secure Information Technology (SIT) - Darmstadt, Germany
Guadalupe Ortiz, University of Cádiz, Spain
Thomas Marx, University of Applied Sciences Bingen, Germany

CENTRIC 2015 Research Institute Liaison Chairs

R. Hariprakash, Arulmighu Meenakashi Amman College of Engineering, India
Muhammad Iftikhar, University of Malaysia Sabah, Malaysia
Kong Cheng, Telcordia Technologies - Piscataway, USA
Gottfried Zimmermann, Stuttgart Media University, Germnay

CENTRIC 2015 Industry/Research Chairs

Eve Maler, Forrester Research, USA
Rainer Falk, Siemens AG - München, Deutschland
Juong-Sik Lee, Nokia Research Center - Palo Alto, USA

CENTRIC 2015 Technical Program Committee

Mohammed A. Akour, Yarmouk University, Jordan
Ana Almeida, Institute of Engineering - Polytechnic of Porto, Portugal
Giner Alor Hernandez, Instituto Tecnologico de Orizaba, México
Rémi Bastide, Université Champollion, France
Carlos Becker Westphall, Federal University of Santa Catarina, Brazil
Stephan Böhm, RheinMain University, Germany
Mahmoud Boufaida, Mentouri University of Constantine, Algeria
Mahmoud Brahimi, University of Msila, Algeria
Lillian Brandão, Universidade Federal de Minas Gerais (UFMG), Brazil
Marco Bruti, Telecom Italia Sparkle S.p.A. - Roma, Italy
Umberto Cerasani, Sophia Antipolis, France
Feng Cheng, Hasso-Plattner-Institute at University of Potsdam, Germany
Kong Cheng, Applied Communication Sciences, USA

Michele Risi, University of Salerno, Italy
Joel Rodrigues, University of Beira Interior, Portugal
Carmen Santero, CNR-ISTI, Italy
Marcello Sarini, Università degli Studi Milano-Bicocca - Milano, Italy
Narasimha K. Shashidhar, Sam Houston State University, USA
Giuseppe Scanniello, University of Basilicata, Italy
Dominique L. Scapin INRIA, France
Guillaume Schaff, itrust consulting, Luxembourg
Peter Schartner, Universität Klagenfurt, Austria
Rainer Schmidt, Munich University of Applied Sciences, Germany
Monica M. L. Sebillo, University of Salerno, Italy
Jacques Simonin, Institut Télécom / Télécom Bretagne, France
Weilian Su, Naval Postgraduate School - Monterey, USA
Roman Y. Shtykh, CyberAgent, Inc., Japan
Hung-Min Sun, National Tsing Hua University, Taiwan
Ruck Thawonmas, Ritsumeikan University, Japan
Genoveffa Tortora, University of Salerno, Italy
Giuliana Vitiello, Università degli studi di Salerno, Italy
Florian Volk, Technische Universität Darmstadt, Germany
Konstantinos Votis, Information Technologies Institute, Greece
Nan Wang, Linkedin.com, USA
Benjamin Weyers, University of Duisburg-Essen, Germany
Sherali Zeadally, University of Kentucky, USA
Haifeng (Kevin) Zhao, University of California - Davis, USA
Gottfried Zimmermann, Stuttgart Media University, Germany

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Control Method of Power Seat Motor for Comfort of Automobile Driver

Myungjin Chung
Department of Mechatronics Engineering
Korea Polytechnic University
Siheung, Republic of Korea
e-mail: mjchung@kpu.ac.kr

Sangyeon Hwang
Department of Research and Development
TSA Co., Ltd.
Siheung, Republic of Korea
e-mail: sangyeonh@hanmail.net

***Abstract*—Automobile seats must provide comfort to drivers and passengers while driving. Drivers should feel comfortable with the seat that has the smooth motion. This can be acquired through the control of seat motor. In this paper, a control method profiling start and stop current is proposed. The soft-start current profile is achieved by increasing the voltage to the motor. With the soft-start current profile, the slow-start motion is possible. The experimental test shows that the proposed control method reduced the reflection sound pressure by 14.3% and the shock acceleration by 85.7%, respectively.**

Keywords-control; power seat; automobile; soft-start; soft-stop; current profile.

I. INTRODUCTION

The seat of an automobile must support the comfort to the driver and passenger while driving. The control method of the seat position is changed from manual type to power type, which means using the motor to increase the comfort of the driver as proposed by Lahiry [1] and Kovener [2]. Lahiry and Kovener used the motor to move the seat from one position to the other position. The seat using the motor is called power seat and that is classified as two-way or four-way power seat according to a number of control direction. By using the motor, several problems, such as vibration, noise, and overcurrent, appear. Researches for the evaluation of these problems are conducted. Cho [3] visualized the automotive power seat slide motor noise, Han [4] suggested the method for automatic measurement of noise and vibration for power seat DC motor in the vehicle, and Seo [5] evaluated the vibration and sound quality according to battery voltage change while moving a power seat in an automobile forward or backward. Those researches suggested the method for measurement and analysis of the noise and vibration generated by the power seat DC motor, which is controlled by step voltage to move the seat. In this study, the method for suppression of the noise and vibration is proposed. The driver should feel comfortable with the seat that has smooth motion, which can be achieved through the proposed method.

In Section II, configuration of the power seat is presented. Then, the control method of the motor for generation of the current profile, such as soft-start and soft-stop, is proposed in Section III. DC motor driver having the proposed control function is developed and experimental test is conducted with the DC motor and the power seat in Section IV.

II. CONFIGURATION OF POWER SEAT

The power seat is classified as two-way power seat or four-way power seat according to the number of control direction. The power seat consists of seat body, motor, motor driver, and actuating mechanism. The motor used in the power seat is classified as recline motor conducting the rotating function of seat back, tilt motor conducting the tilting function of seat base, slide motor conducting the moving function of seat base horizontally, and height motor conducting the moving function of seat base vertically according to function as shown in Figure 1. The DC motor, having properties such as high starting torque, good output efficiency for input current, and easy rotating control, is used for the power seat. The motor driver adopts the electric driving type using the electric switching device.

III. CONTROL METHOD

The driver should feel comfortable with the seat that has smooth motion, which can be acquired through DC motor control having soft-start and soft-stop current profile as shown in Figure 2. In this figure, current profile by proposed control method is smooth compared to current profile due to step input voltage. By the adoption of soft-start and soft-stop current profile to DC motor control, the power seat can be moved smoothly.

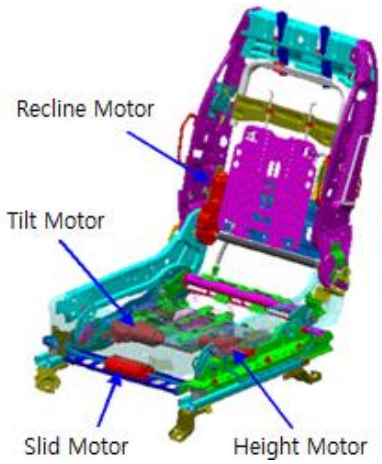


Figure 1. Configuratin of power seat having recline motor, tilt motor, slide motor, and height motor.

 ISBN: 978-1-61208-440-4

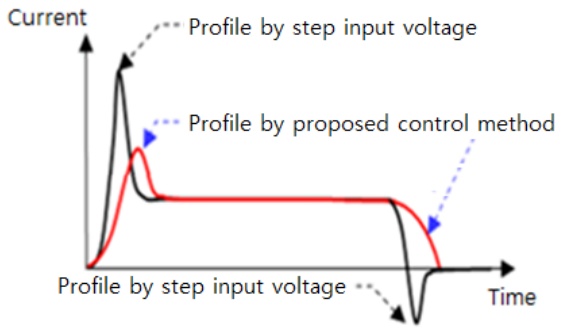


Figure 2. DC motor control method having soft-start and soft-stop current profile.

The soft-start current profile can be achieved by increasing the voltage applied to the DC motor as shown in Figure 3. The shape of the soft-start current profile can be controlled according to increasing time (Ti), which is changed according to the friction force of guide rail in the power seat. The control block diagram of DC motor driver, which is using the electric switching device, is shown in Figure 4. The gate signal is generated by controller according to current flowing to the DC motor. The soft-start current profile is controlled by the gate signal. The increasing time (Ti) for soft-start and decreasing time (Td) for soft-stop can be controlled in the microprocessor.

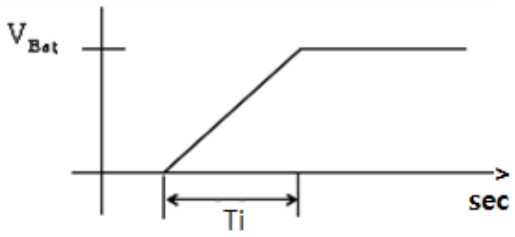


Figure 3. Applying voltage to DC motor for the generation of the soft-start current profile.

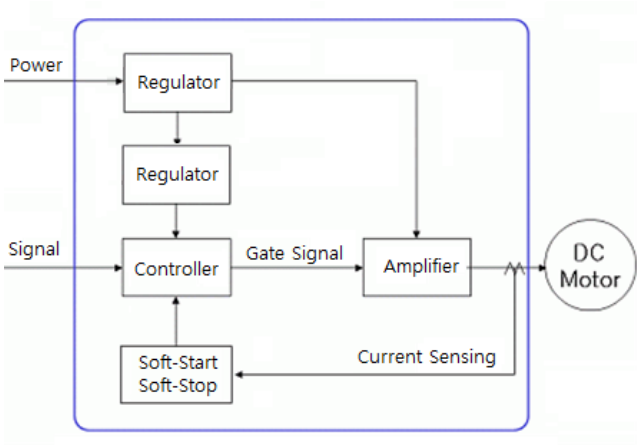


Figure 4. Control block diagram of DC motor driver.

IV. Experimental Results and Discussion

DC motor driver having functions shown in Figure 4 is developed as shown in Figure 5. The developed motor driver circuit consisted of control circuit for control of motor by microprocessor and power circuit for generating the current. Figure 6 shows the experimental setup for the performance test of developed DC motor driver. In this setup, power supply, current probe, and digital oscilloscope are used to measure the driving voltage, current, soft-start time, and soft-stop time.

Figure 7 (a) shows the measured motor current profile without control of voltage to DC motor in case of motor start. The seat has the sudden-start motion through the current profile shown in Figure 7 (a). The driver may feel discomfort from the sudden-start motion of the seat. Figure 7 (b) shows the measured motor current profile with control of voltage to DC motor in case of motor start. The soft-start time is measured as 0.226sec, and this time can be controlled in the microprocessor. The seat has the slow-start motion through the current profile shown in Figure 7 (b). The driver may feel comfort from the slow-start motion of the seat. Figure 8 shows the measured motor control signal and current profile generated by the developed motor driver in case of soft-start and soft-stop.

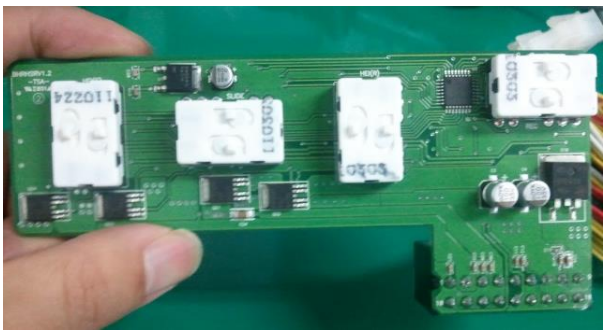
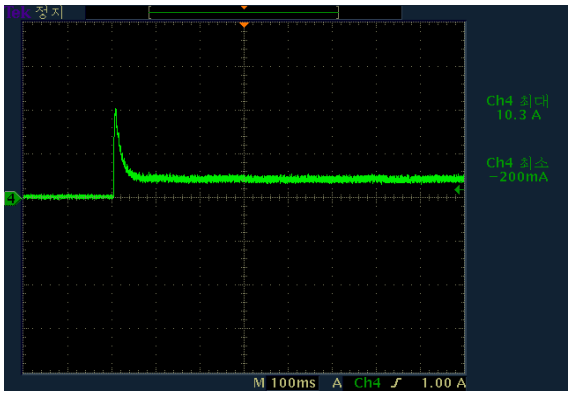
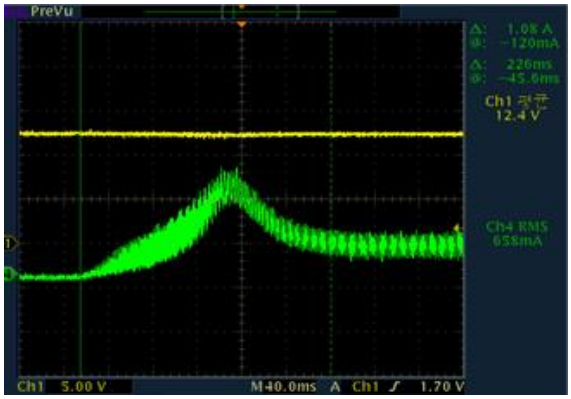


Figure 5. Developed DC motor driver having the functions of soft-start and soft-stop.

Figure 6. Experimental setup for the performance test of developed DC motor driver.



(a)



(b)

Figure 7. Measured motor current profile (a) without control of voltage and (b) with control of voltage to DC motor.

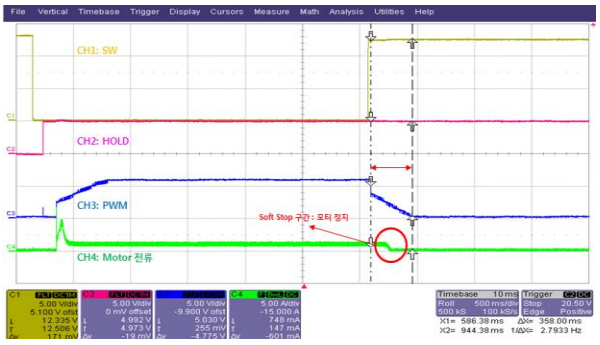


Figure 8. Measured motor control signal and current profile generated by control of voltage to DC motor with the developed motor driver.

Table I lists the performance of the developed motor driver. The soft-start and stop-time can be controlled from 0.02sec to 2sec. This time duration is determined according to properties such as friction force between guide rail of seat.

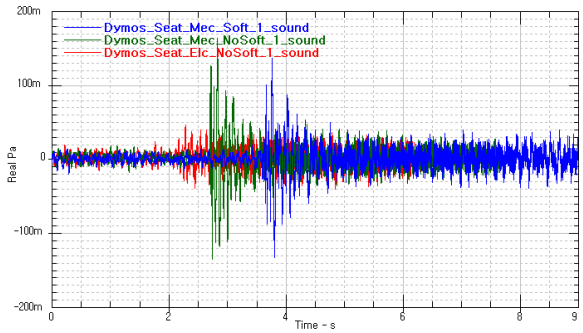
Figure 9 shows the experimental setup for the performance test of seat motion by applying the proposed control method in the dead sound chamber. Pressure meter is used to measure the reflection sound pressure and the accelerometer is used to measure the shock acceleration.

Figure 10 shows the measured reflection sound pressure and shock acceleration according to the absence and presence of soft-start and soft-stop control method.

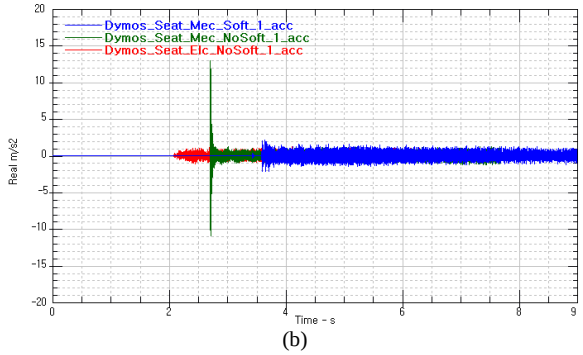
TABLE I. PERFORMANCE OF DEVELOPED DC MOTOR DRIVER

Parameter	Value	Unit
Driving voltage	6-24	V
Driving current	20	A
Soft start time	0.02-2	sec
Soft stop time	0.02-2	sec

Figure 9. Experimental setup for the performance test of seat motion by applying the proposed control method in the dead sound chamber.



(a)



(b)

Figure 10. Measured (a) reflection sound pressure and (b) shock acceleration.

In Figure 10 (a), the reflection sound pressure of non-controlled motor is 160mPa and the reflection sound pressure of controlled motor is 140mPa. In Figure 10 (b), the shock acceleration of non-controlled motor is 14m/sec^2 and the shock acceleration of controlled motor is 2m/sec^2. From the verification test, the reflection sound pressure is reduced by 14.3% and the shock acceleration is reduced by 85.7% by applying the proposed control method.

V. CONCLUSIONS

In this paper, the control method of the motor for the generation of current profile, such as soft-start and soft-stop, is proposed. The soft-start and soft-stop current profiles are achieved through the voltage control applied to the DC motor. DC motor driver is developed, and experimental test using the developed motor driver is conducted to verify the proposed control method. From the verification test, the reflection sound pressure is reduced by 14.3% and the shock acceleration is reduced by 85.7% by applying the proposed control method. From the study, we expect that the driver can experience comfort upon applying the proposed control method.

In the future, we will evaluate the comfort level of drivers, when the control method is applied to an automobile.

ACKNOWLEDGMENT

Following are results of a study on the "Leaders Industry-university Cooperation" Project, supported by the Ministry of Education, Science & Technology (MEST).

REFERENCES

[1] A. Lahiry, S. Chanana, and S. Kwnwar, "Automation in automobiles: power seat," The Second International Conference on Machine Vision (ICMV 2009) IEEE, Dec. 2009, pp. 158-160.

[2] R. W. Kovener, "Power seat with memory," International Journal of Vehicle Design, vol. 2, pp. 111-117, 1981.

[3] Y. T. Cho and J. S. Bolton, "Visualization of automotive power seat slide motor noise," National Conference on Noise control engineering (Noise-Con 14) INCE, Sep. 2014, pp. 181–188.

[4] H. S. Han, B. W. Jeong, G. H. Kim, and D. H. Song, "Automatic measurement of noise and vibration for power seat DC motor in the vehicle," The Spring Conference on The Korean Society for Noise and Vibration Engineering (SCKSNVE 2002) KSNVE, May 2002, pp. 1142-1147.

[5] K. S. Seo, D. S. Choi, and K. S. Kim, "Evaluation of vibration and sound quality according to battery voltage change while moving a power seat in an automobile forward or backward," Journal of Central South University, vol. 21, pp. 3844-3849, 2014.

 ISBN: 978-1-61208-440-4

Taking Printed Books into Internet of Things

Harri Ketamo
Satakunta University of Applied Sciences
Pori, Finland
harri.ketamo@samk.fi

***Abstract*—The aim of this paper is to describe a development process for bookAI, a system that takes printed books into the Internet of Things by applying smartphones, different identification methods, mobile internet, and cloud computing. BookAI connects the books into the Internet with NFC, RFID or barcode identifiers that are read by smart phones. The challenge in the book identification is recognizing the object from various coding sets and connect it to the existing metadata and content: In bookstores the EAN-13 barcodes are used, while in libraries there are several RFID standards as well as several barcode formats in use. Metadata and available digital content, related to the physical object, are read from different sources. The received data is not in a structured format, which means there are incomplete records and missing data that makes challenges for further computing. BookAI applies semantic computing to build a network of detailed relations between the books and personal reading interests. The personal reading interests are recorded as self-evaluations of the users, so bookAI requires several evaluations before it can make personalized suggestions. After the system has learned the user preferences, it enables searching interesting books from the libraries, bookstores or even from private book shelfs by optimizing the personal content and context map taught by the user. One of the challenges, in designing and end-user product, is user interface. The first version of the bookAI was evaluated complex to use, so the second revision is built based on end-user feedback. Computationally built semantics and ontologies save remarkably time when modeling big domains or big data like library collections. The future research is focusing in building methods to collect more complete datasets about books as well as designing the overall user experience of the system.**

Keywords; smartphones; mobile internet; books; libraries; cloud computing

I. Introduction

The Internet of Things (IoT) is the network of physical objects embedded with electronics, software, data and connectivity. In IoT, each thing is uniquely identifiable and is able to interoperate within the existing Internet infrastructure. This has been expected to extend the understanding of the physical world, leading to innovative services and increase in efficiency and productivity [1][2].

The relation between books and internet, as well as smart museums are widely studied [3]. However, in the discussions of IoT applications printed books are mentioned only in relatively small number of studies [4][5]. In fact, in terms of information retrieval, we can say there is no big difference between real-world pages or books compared to web pages or social media.

Social media services, such as YouTube and Flickr, contain enormous number of content related to each other. Search engines can list numerous pieces of content that matches more or less perfectly to keywords. A common method to increase information accessibility in social media applications is tagging. However, when tags are used only as single words, we easily end up to information overload.

Furthermore, in social media, we do not have standardized way to tag content. In fact, tagging the content in an optimal way is a difficult task for several reasons. Cultural background, educational background, community and its social behavior, as well as context where tagging is constructed affects enormously the selection of tags. The term ‘context’ can be understood in many ways. In this study, context is understood to cover all conditions, physical, social and mental, which can be considered as causes or consequences of activity.

The structure of the paper is following: in Section 2, related work and background information is presented. In Section 3, research objectives are described. In Section 4, the research results are presented and explained. In Section 5, conclusions are discussed.

II. Related Work

In this study, semantic networks based approach is used to model a user's personal interests in order to optimize the recommendations the user gets. In this approach, the originality and novelty is in getting a detailed understanding of what a users likes, prefers or needs to know in order to reorganize the content map and recommendations to meet the needs of the user. In other words, this is more related to user centric computing than information retrieval. Author’s previous studies on the subject has been made especially related to learning sciences [6] and information overload [7] [8]. The fours on this section is on describing the known challenges from existing research: 1) Challenges in tagging the content, where also finding the most important words from the texts is considered as tagging, 2) challenges in adaptive media, 3) challenges in behavior modeling and 4) known work in connecting media objects to digital systems.

Tagging is very subjective and numerous research is done in order to improve user experiences and information retrieval in social media [9][10][11]. Unclear, or in worst case misleading, tagging leads to information loss in social

 ISBN: 978-1-61208-440-4

media. Especially, constructing storytelling or narration between pieces of user generated content becomes impossible without socially constructed tagging semantics. Furthermore, tagging can be seen as a one key element when building platforms for personalized services, but understanding semantics brings new dimension to text analytics. Recently, Google had added semantics into its searches, but many educational subjects require more detailed conceptual models for successful content personalization. Furthermore, there are other personalization and adaptation solutions for media [12].

Adaptive and/or personalized media is always designed to produce optimized user experiences. In adaptation, the mechanics optimize the content with technologies that can be divided into two main groups: indirect (static) adaptation and direct (dynamic) adaptation. In indirect adaptation, the rules are fixed beforehand by developers. Indirect adaptation is based on statistical rules, decision trees, state machines, or the cumulative effects of several fixed functions. In dynamic adaptation, the mechanics track the user and optimize the content according to a user's behavior. In other words, dynamic adaptation is based on machine learning. Dynamic adaptation requires at the very least 1) a user model, 2) a context model and 3) artificial intelligence [13][14][15].

Semantic networks, also known as conceptual graphs, are knowledge representations constructed with directed or undirected graphs [16][17]. Semantic neural networks (SNN) are generally used for processing natural languages [18]. However, SNNs, as knowledge representations are relatively extensible and they have been used, for example, to model mental disturbances [19]. On the other hand SNN can be utilized to model the characteristics of users, profiles, patterns of behavior, and skill levels in order to support or challenge the performance of individuals.

Parallel methods, such as behavior recording [20][21] and behaviour mining [22][23] have been studied and used in the game industry for some time. Behavior recording refers to game development and behavior mining usually refers to intrusion detection in networks, etc. Because the idea of adaptive educational systems is to produce individual and optimized learning experiences [24][25], high end user models, as well as methods, are relatively complex. In the high end solutions intelligence is based on neural, semantic, or Bayesian networks, as well as genetic algorithms [26][27] [28].

Applying NFC and RFID in end-user products has been studied e.g., in social mobile games with educational dimensions, with shared social experience and physical interaction between players [29][30][31]. Furthermore, location based NFC games and location aware NFC UIs allow users to play games in mixed reality in that they can interact with both real and virtual objects within that location [32][33][34][35]. In these environment NFC and RFID are used to extend mobile devices as media controllers, but not only because of the object. The key idea has been to connect the textual content of the physical objects into digital systems.

III. Research objectives

The aim of this paper is to describe a development process for bookAI. The technological challenge of this study is related to the complexity of the system: i) Recognizing a book with NFC, various RFID and barcode formats, 2) collecting unstructured and incomplete information from various sources, 3) modeling context maps as well as user's personalization maps based on incomplete data and 4) visualizing complex phenomena in a way that end-user can find it useful.

This is a design study including i) algorithm and system design, ii) experimental tests with real world big data and iii) piloting study with small number of real-world users.

The first version of the system was evaluated with small (n=10) end-user group that consist of Finnish librarians. The aim of the pilot study was to proof that the concept works. Test persons reported only bugs and challenges in usability. Furthermore, they were asked to give ideas and improvement comments for next prototype that should be piloted with library customers in Fall 2015. No survey data was collected after piloting.

When the content consists of millions of objects and billions of relations, compared to previous studies with only thousands of objects and hundred millions relations [25][36], the computational management of the relations between the objects becomes a challenge in terms of smooth user experience. When redefining the relations between content objects, this system can be managed through semantic networks in a computationally efficient way. The data-sample size is more than 2 million books from Satakunta regional libraries. This sample data contains with more than 500GB of descriptive data about the books, and the data enables to build more than 3 billion connections between books. This requires also system design from Big Data point of view.

The challenge in the book identification is recognizing the object from various coding sets and connect it to the available metadata and content: In bookstores the EAN-13 barcodes are used, while in libraries there are NFC, several RFID standards as well as several barcode formats in use. Metadata and available content (e.g., abstracts) are read from different sources. The received data is not is a structured format, which means there are incomplete records and missing data that makes challenges for computing.

IV. Results

Results section is presented in the order required to start to use the system: First the context model is explained, secondly how the system learns the user preferences and finally how it adopts to user's needs. The results and feedback from the pilot study is presented with the topic it belongs to.

Context model design starts with defining the dependencies (or proximity) between single objects (books). In this study, the proximity is defined by name of the book, keywords and other metadata given by book publisher/librarians and abstract text. All this information is

 ISBN: 978-1-61208-440-4

not necessary available: metadata is subjective, It varies between different libraries and bookstores. Also, there are differences in abstract texts and surprisingly also with book names. The differences in book names is related to age of records. Some of the oldest records we use was form eighties, and so there were misspellings, additional info in name-record, etc.

The proximity is calculated by searching for similar words in similar sections (name, keywords, abstract) and weighting the findings according to frequencies of the words. An example of the idea is visualized in Figure 1.

The upper book (Birds of Europe) do have tags/keywords (Europe, biology, geography,…) and an abstract text. The centre book (Capital cities Europe) do have a same word in the name (Europe), two same keywords (Europe, geography) and two same words in the abstract. When summarizing the total proximity value, keywords are as valuable as name and abstract together. So the total score in this example is 7 hits. In the final algorithm, the words are also weighted based on their reversed frequency, which means words like "news" or "sports" do have less weight than words like "Oslo" or "Owl", so this also effects the final scoring, but is excluded from this example in order to keep this simple.

The book in the bottom (Capital in the twenty-first century) has one same word in name that Capital cities Europe (Capital) and the same word (capital) in abstract. The total proximity score for the book is two, even though we know there is relatively little common between capital cities and money. That's why there is always a minimum value for proximity that has to be passed before the proximity has been validated. In this case (without word weighting) it would be 4 or 5 hit points. In other words, couple of common words is not enough to show proximity in this algorithm.

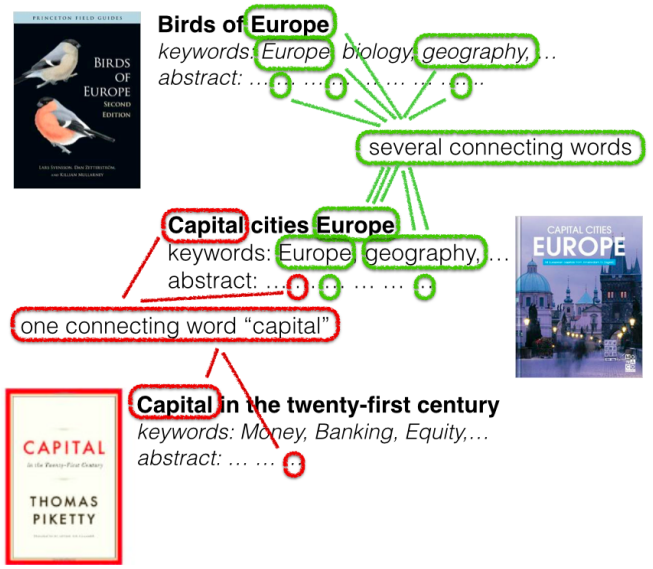


Figure 1. Calculating the distance between the objects

In the example, there is no common words between "Birds of Europe" and "Capital in the twenty-first century", which is relatively obvious. Naturally, pronouns, numerals, participles, etc. are excluded from the meaningful word set. The biggest difference in this study is that the non-meaningful words are excluded computationally. If a word occurs frequently in texts, the word is useless from text prediction point of view. Rare words, however, are useful but alone they might be misleading, like previous example on word "capital". That's why weighting the words as well as requiring several common words seems to give reasonable results as a full computational approach.

The biggest challenge in building proximity map between the books is in the total number of relations between the books. For example (Figure 2, in Finnish) the 10 nearest book to "Birds of Europe" forms a network of 46 relations when maximum is 50 relations between 10 books. In other words, the biggest challenge is not to find proximate books, but rank the nature of the proximity and manage the big data.

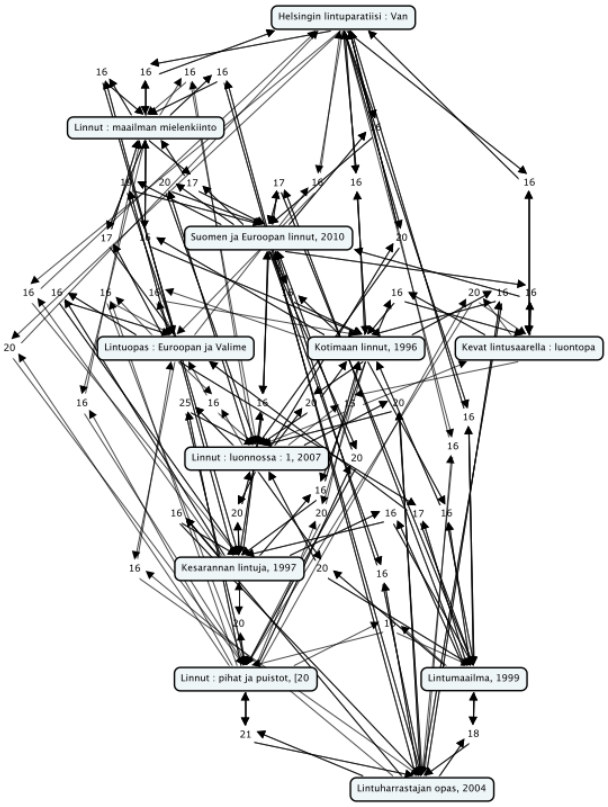


Figure 2. Example on sub-network of books in neighborhood of "Birds of Europe"

Based on the data from bookAI, there are approximately 3 000 000 000 two-way relations between 2 000 000 books, which means, on average, that each book does have approximately 750 single relations to other books. Naturally, when just looking for the most proximate books, we don't need two-way relations for all these thousands of books: rank

 ISBN: 978-1-61208-440-4

ordered index is enough for defining the data and at the same time, performance is fast.

In other words, when building a end-user version of this solution, we won't use the full computational model for all the calculations. Most of the information presented in end-user user interface is pre-computed indexes which ensures the smooth performance of the system even with high number of users.

Another challenge is visualize complex data. N-dimensional networks are not interesting for most of the users, so the information should be presented in more understandable form.

In our prototype application, we have transformed n-dimensional network into 2-dimensional grid/heatmap (Figure 3, in Finnish). The idea of the grid is that in the the most proximate books are always neighboring and the heat (from red to yellow) shows the proximity to the requested books (Birds of Europe in this case). In other words, we flatten the information in some sense, but increase the instant readability of the data.

Figure 3. Example visualization on two-dimensional network neighborhood of "Birds of Europe" (names in Finnish).

The texts in Figure 3 are in Finnish, including the name of the book, author of the book, ISBN and language and location of the book. The heatmap behind the text is the most important factor when visualizing how proximate the books are. What closer to red the background is, that proximate the books are. If the background is white (no heat) there is not much common between the book in whit background and the center book (the book we are focusing now).

One major challenge with this data visualization is, that we loose some books: if a book is 9th proximate to starting book and also 9th proximate (or less) to it's successors, it will be dropped to the grid even though there will be books on the grid that has very little to do with the starting book. This challenge will be fixed in future studies, when we'll bring different type of visualizations into end-user testing and so we can empirically find what kind of visualizations users really like.

The context model is a general model about the context we're dealing with, it is built with unsupervised machine learning model, where the learning is based on given dataset. The personalization / adaptive model, described next, is built over context model by applying supervised learning -type of methods, where learning is based on input given by the user.

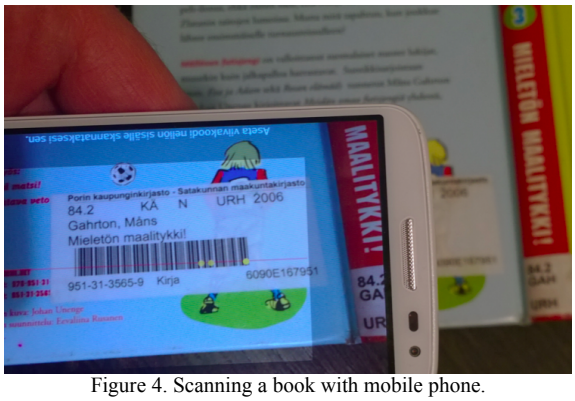


Figure 4. Scanning a book with mobile phone.

In end-user application, the process starts by recognizing the book (Figure 4) we are working with. The recognition of the book is done by scanning the barcode behind the book (either library dependent identifier or global EAN barcode) with which we can find the ISBN of the book. The scanning is done with mobile device's camera. The mobile device is the connecting object between books and the internet.

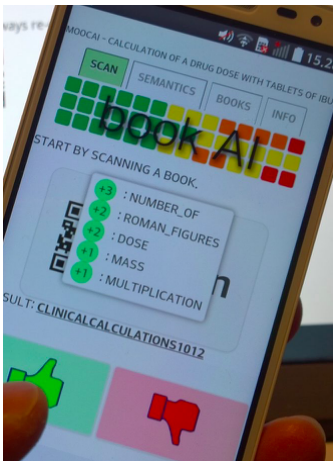


Figure 5. Main interface of the first version of the system and evaluation of a scanned book.

ISBN number is the key element of the context model, while in user model side words are the key elements. This is crucial in order to e.g., deal with multiple ISBNs related to

same book. In fact, is is common that there are tens of different ISBNs related to the same specific book because of e.g., new editions, which automatically get a ISBN number.

After the book and ISBN is recognized user can evaluate the book by giving thumbs up or thumbs down for a book (Figure 5). In this example we are evaluating a book about medical treatment, so when giving thumbs up (user likes the book or get the information he/she wanted), the application shows the words that get additional positive weight from the user.

Each evaluation (thumbs up / thumbs down) increases the information about user's preferences. In other words, algorithm learns user preferences case by case and improves the explanative power of user model case by case.

In the first version of the bookAI, the scanning button and evaluation was placed in the front of the main screen. Users were expected to either scan a new book, evaluate a previous one or open a new tab. The proximity maps was placed on the next tabs, which user should open in order to see the scanned maps and personalizations. The user interface was found too complex to use and test users reported that is was difficult to recognize which book they are evaluating or what proximity map they were following.

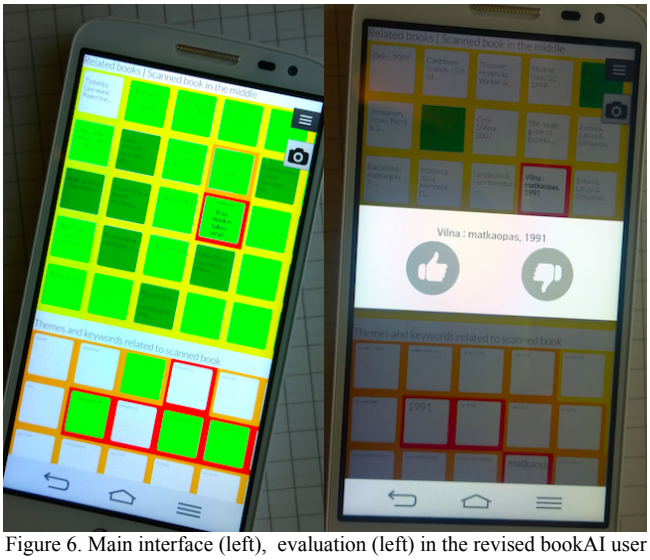


Figure 6. Main interface (left), evaluation (left) in the revised bookAI user interface.

This map can be seen as personalized reading assistant showing always new books to read. From user experience point of view, it is important to show also books that user may not like. If showing only books that are probably liked, user will sooner or later miss all the interesting new books that are close to liked books but with vocabulary not exactly the same.

In the revised version of the bookAI (Figure 6) the user interface was rebuilt in order to override all these reported challenges. The main screen in the heat map with smaller grid, consisting only 30 most proximate books, followed by semantic heat map, showing the words and themes connected to scanned book. In the top-right corner of UI there is two icons: upper icon is for menu, from which user can see the list of canned books. The list is ordered from newest to oldest, so user can immediately see what heatmap he/she is following. The camera button takes to scanner and evaluation is immediate action after scan. User can override the evaluation any time, but in this we we ensure user knows for sure what book he/she is evaluating.

In Figure 7, an example about fictive reader's opinions about bird-books: User has evaluated altogether 10 bird books, from which 4 can be seen in Figure 6. Books covering natural environments (or environment is not mentioned) are thumbed up and books related to built environment are thumbed down.

In Figure 7, the background heat map (context map) and the texts are the same as in Figure 3. The personal layer consist of a) books with blue background, representing the probability that user don't like the book, b) green backgrounds representing that user most probably like the books (what lighter green that higher probability) and c) grey backgrounds showing that user data about the book is noisy.

About the personalized map user can immediately see which books he/she likes and how the books are related to centre book.

Figure 7. Example on two-dimensional personalized network neighborhood of "Birds of Europe" (names in Finnish)

From machine learning point of view it is crucial to give extensive evaluations. If the system starts to get only positive evaluations from very narrow context and the context would get even narrower all the time, there is soon nothing to learn any more. From information theory point of view the system needs entropy in order to learn more.

The recommendation system can be used outside the library e.g., in online education (Figure 8) as well as for presenting additional information already online (Figure 9).

When the online content is added into bookAI as a 'book', we can connect physical books in the library into the digital course (Figure 8). The idea in case of online learning is that

user thumbs up content that he/she understands and thumbs down the content he/she don't understand. The application, in this case, shows the books related to online content and also predicts if user understands the specific topic of the book.

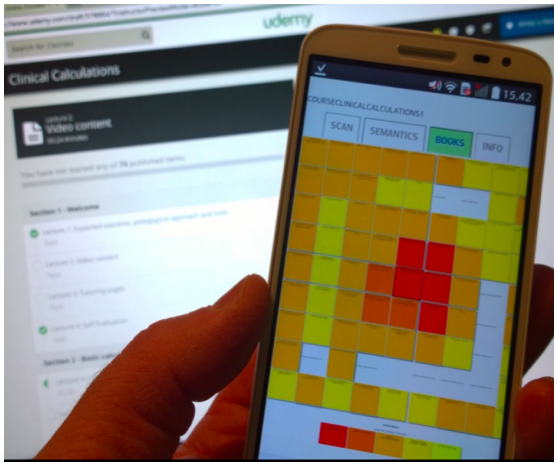


Figure 8. Connecting library to online course

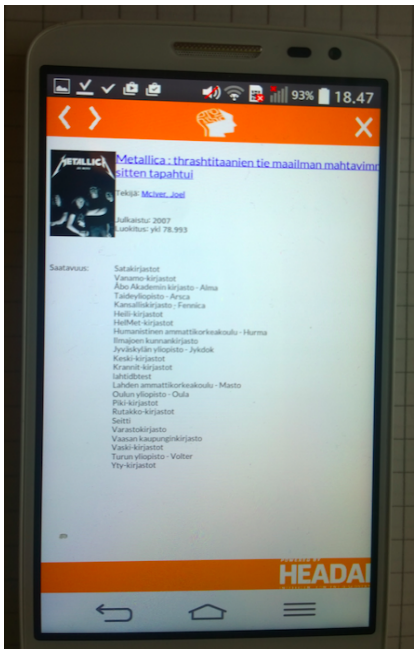


Figure 9. Connecting libraries existing online content to books.

The same goes with e.g., household appliances: When scanning the barcode of microwave oven, a user will get a heat map about books related to cooking, microwave technology and ovens. The personal layer depends on user's personal preferences (cooking or technology or both). In all cases, the books will be connected to other objects via user's personal mobile device.

Furthermore, existing online content about the books, can be connected to bookAI. With this feature, user can immediately find additional information about the book without logging into library's information system. The additional content is can be accessed directly from heatmaps.

V. Conclusions

The aim of this paper was to describe the production and design challenges of bookAI. The bookAI maps the content automatically, algorithm based, which means no manually designed semantics or ontologies are needed. Computationally built semantics and ontologies save remarkably time when modeling big domains or big data like library collections. This kind of non-supervised learning requires, however, a large learning data and the idea is not applicable for small or medium size learning data. For example, single web page is far to small on content in order to be mapped, but conference proceedings with hundreds of papers may be enough.

When comparing bookAI to more known literature recommendation systems, such as e.g., Amazon Bookstore's crowd sourced 'the buyers of this book also bought the following books' -type of recommendation systems, we can immediately point out two major benefits.

First of all, when crowd sourcing recommendations, we need a big user data in advance. Without big user data, we can give suggestions for only a small number of book. The minimum requirement for such data would be equal to the number of books and in that case there should be exactly two bought books per user and no overlapping at all. Naturally this never exists in real world and the required data would be remarkably bigger. With bookAI we can start recommendations immediately and the recommendations get more exact during the use of the system.

Secondly, when recommendations are based on previous buys, the first buys starts to control the next buys and finally there is no room for real, empirically grounded, recommendations. In other words, the crowd sourced recommendation data is corrupted by the meas how it was collected. BookAI can bring recommendations based on the semantic proximity of the books and empirically collected user preferences. This gives the user better understanding on what is available without limitations of somewhat misleading way to use empirical data.

BookAI learn user skills and competences based on self-evaluations (thumbs up / down) and connects preferences to general context map. Using machine learning for personalization enables developer to bring any content to system. Pre-fixed preference lists, defined by users, are always to restricted and out of date. The downside on using supervised machine learning is that it requires tens of evaluations before the system has learned something about the user. When the user has spent some time to do evaluations (teaching the system), the user model is transferrable to other contexts as well, so in long run, the time it takes to start teaching the system pays off.

The visualization of the data is designed to be easily understandable. This, however, requires more research. According to user feedback received from librarians, the data

 ISBN: 978-1-61208-440-4

is readable but maybe still a bit too complex for average library user. This indicates that more research has to be done in visualization side in order to build a successful consumer platform.

References

[1] D. Bandyopadhyay and J. Sen. "Internet of Things: Applications and Challenges in Technology and Standardization". Wireless Personal Communications, Vol 58(1), 2011, pp. 49-69.

[2] D. Miorandia, S. Sicarib, F. De Pellegrinia, and I. Chlamtaca. "Internet of things: Vision, applications and research challenges". Ad Hoc Networks, Vol 10(7), 2012, pp.1497–1516.

[3] L. Atzori, A. Iera, and G. Morabito. "The Internet of Things: A survey". Computer Networks, Vol 54(15), 2010, pp.2787-2805.

[4] H. Pencea. "Smartphones, Smart Objects, and Augmented Reality". The Reference Librarian, Vol 52(1-2), 2010, pp.136-145.

[5] A. Koivisto, H. Ketamo, and E. Hammais. "Combining the Dimensions of Written and Digital Media in a NFC-based Non-linear Adventure Game for Children". In proceedings of IFLA WLIC 2014. August 16-22, 2014, Lyon, France.

[6] H. Ketamo. "Learning Fingerprint: Adaptive Tutoring for MOOCs". In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications. Jun 23, 2014, Tampere, Finland, pp. 2349–2358.

[7] H. Ketamo. "Sharing Behaviors in Games and Social Media". International Journal of Applied Mathematics and Informatics, vol. 5(1), 2011, pp. 224-232.

[8] H. Ketamo. "Managing Information Overload - Teachable Media Agents". In Proceedings of the 8th International Conference on Intellectual Capital, Knowledge Management & Organisational Learning – ICICKM 2011. 27-28 October 2011, Bangkok, Thailand, pp. 301-308.

[9] E. Agichtein, C. Castillo, D. Donato, A. Gionis, and G. Mishen. "Finding high-quality content in social media". In Proceedings of the International Conference on Web Search and Web Data Mining, WSDM'08. Palo Alto, California, USA, February 11-12, 2008, pp. 183-194.

[10] P. Heymann, G. Koutrika, and H. Garcia-Molina. "Can social bookmarking improve web search?" In Proceedings of the International Conference on Web Search and Web Data Mining, WSDM'08. Palo Alto, California, USA, February 11-12, 2008, pp. 195-206.

[11] B. Sigurbjörnsson and R. van Zwol. "Flickr tag recommendation based on collective knowledge." In Proceedings of the 17th International Conference on World Wide Web, WWW'08. Beijing, China. April 21-25, 2008, pp. 327-336.

[12] J. Ahn, P. Brusilovsky, D. He, J. Grady, and Q. Li. "Personalized web exploration with task models." In Proceedings of the 17th International Conference on World Wide Web, WWW'08. Beijing, China. April 21-25, 2008, pp. 1-10.

[13] J. Manslow. "Learning and Adaptation". In Rabin, S. (ed.) AI Game Programming Wisdom. Massachusetts: Charles River Media, Inc., 2002, p. 557-566.

[14] P. Brusilovsky. "Adaptive Hypermedia". User Modeling and User-Adapted Interaction. vol 11, 2001, pp. 87-110.

[15] J. Eklund and P. Brusilovsky. "InterBook: an Adaptive Tutoring System". UniServe Science News, vol 12, 1999.

[16] J. F. Sowa. "Semantic networks". In Shapiro, S.C.(ed.) Encyclopedia of Artificial Intelligence, Wiley, New York, 1987.

[17] J. F. Sowa. "Conceptual Graphs". In van Harmelen, F., Lifschitz, V., & Porter, B. (eds) Handbook of Knowledge Representation, Elsevier, 2008, pp. 213-237.

[18] D. E. Shuklin. "The Structure of a Semantic Neural Network Realizing Morphological and Syntactic Analysis of a Text". Journal Cybernetics and Systems Analysis, vol 37,5, 2001, pp.770-776.

[19] A. B. Geva and A. Peled. "Simulation of cognitive disturbances by a dynamic threshold semantic neural network". Journal of the International Neuropsychological Society, Vol 6(5), 2000, pp. 608 - 619.

[20] M. Bowling, J. Furnkranz, T. Graepel, and R. Musick. "Machine learning and games". Machine Learning, vol 63, 2006, pp. 211-215.

[21] R. Houlette. "Player Modeling for Adaptive Games". In Rabin, S. (ed.) AI Game Programming Wisdom II. Massachusetts: Charles River Media, 2003, pp. 557-567.

[22] S. Mukkamala, D. Xu, and A. H. Sung. "Intrusion Detection Based on Behavior Mining and Machine Learning Techniques". Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2006, pp 619-628.

[23] Y. H. Kuo, Y. M. Huang, J. N. Chen, and Y. L. Jeng. "Extended Real-Time Learning Behavior Mining." In proceedings of Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05), 2005, pp. 440-441.

[24] P. Brusilovsky and C. Peylo. "Adaptive and intelligent Web-based educational systems". International Journal of Artificial Intelligence in Education. vol 13, 2-4, 2003, pp. 159-172.

[25] H. Ketamo. "Balancing adaptive content with agents: Modeling and reproducing group behavior as computational system". In proceedings of 6th International Conference on Web Information Systems and Technologies, WEBIST 2010, 7-10 April 2010, Valencia, Spain, vol 1, pp. 291-296.

[26] J. Reye. "Student Modelling based on Belief Networks". International Journal of Artificial Intelligence in Education, vol 14, 2004, pp. 63-96.

[27] K. M. Kim, J. H. Hong, and S. B. Cho. "A semantic Bayesian network approach to retrieving information with intelligent conversational agents." Information Processing & Management, vol 43,1, 2007, pp. 225-236.

[28] P. J. F. Lucas. "Bayesian network modelling through qualitative patterns". Artificial Intelligence, Vol 163, 2, 2005, pp. 233-263.

[29] P. C. Garrido, G. M. Miraz, I. L. Ruiz, and M. A. Gomez-Nieto. "Use of NFC-based Pervasive Games for Encouraging Learning and Student Motivation". Near Field Communication (NFC), 2011 3rd International Workshop on, IEEE 2011, pp. 32-37.

[30] A. Nandwani, P. Coulton, and R. Edwards. "NFC Mobile Parlor Games Enabling Direct Player to Player Interaction". In proceedings of Near Field Communication (NFC), 2011 3rd International Workshop on Near Field Communication, 22 February 2011, Hagenberg, Austria, pp. 21-25.

[31] A. Koivisto, H. Ketamo, and E. Hammais. "Combining the Dimensions of Written and Digital Media in a NFC-based Non-linear Adventure Game for Children". In proceedings of IFLA WLIC 2014. August 16-22 2014, Lyon, France.

[32] O. Rashid, P. Coulton, R. Edwards, and W. Bamford. "Utilising RFID for mixed reality mobile games." In proceedings of Consumer Electronics, 2006. ICCE '06. 2006 Digest of Technical Papers. International Conference on, IEEE 2006, pp. 459-460.

 ISBN: 978-1-61208-440-4

[33] I. Sánchez, M. Cortés, J. Riekki, and M. Oja. “NFC-based interactive learning environments for children”. In Proceedings of the 10th International Conference on Interaction Design and Children, IDC '11, June 20-23, 2011, Ann Arbor, Michigan, USA, pp. 205-208.

[34] T. Vajk, P. Coulton, B. Bamford, and R. Edwards. “Using a Mobile Phone as a “Wii-like” Controller for Playing Games on a Large Public Display”. International Journal of Computer Games Technology Volume 2008, Article ID 539078, 6 pages.

[35] G. Broll, R. Graebsch, P. Holleis, and M. Wagner. “Touch to play: mobile gaming with dynamic, NFC-based physical user interfaces”. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, MobileHCI '10, September 2010, Lisboa, Portugal, . pp. 459-462.

[36] H. Ketamo. “Self-organizing content management with semantic neural networks”. In Recent Advances in Neural Networks: Proceedings of the 10th WSEAS International Conference on Neural Networks (NN’09), Prague, Czech Republic, 23-25.3. 2009, pp.63-69.

 ISBN: 978-1-61208-440-4

Low Cost Domotic System based on Open Hardware and Software

Valentim Realinho
Polytechnic Institute of Portalegre
Portalegre, Portugal
e-mail: vrealinho@estgp.pt

***Abstract*—With the proliferation of wireless networks and their integration in electrical equipment, the possibilities for monitoring and control of electrical devices increased a lot expecting with this an increased usage of home automation systems. Nowadays, there are in the market several manufacturers that provide fairly comprehensive domotic solutions. However, these solutions are proprietary and have very high costs making it difficult to adopt. In this work, it is presented a low cost home automation system based on open hardware and software with support for a wide range of equipment. The system is based on a Raspberry Pi network and can be easily controlled via a mobile device. It is also possible to configure alerts that notify users when certain events occur and monitor the power consumption.**

Keywords-Domotic; home automation; smart home; mobile device; Raspberry Pi; low cost; open.

I. INTRODUCTION

The word domotic is frequently associated with the concept of smart home technology while allowing the management of electrical devices in a space in order to increase comfort, energy efficiency or safety. This word is the result of the junction of the word domus (house) with telematics (telecommunications and computing), assuming these concepts the existence of computer systems that monitor and control the set of existing electrical equipment. Despite home automation systems exist for more than 30 years, its adoption has been rather slow. Among the main barriers for this low adoption, stands out: (i) the high cost, (ii) the lack of flexibility to interconnect multi-vendor equipment, (iii) management skills and limited or difficult configuration, and (iv) the difficulty in ensuring security mechanisms [1].

Many of the facilities are made by users themselves [1] with DIY (Do It Yourself) equipment, and correspond in most cases to simple systems that enable remote control of equipment, task scheduling or security systems. Other equipment or systems imply the need for structural changes to be installed and its installation by skilled professionals increasing therefore costs.

There are plenty of cases of use of home automation systems. Lighting systems can be programmed to turn off under certain conditions or to automatically turn on when a person enters a dimly lit room. Shutters, HVAC (Heating, Ventilation, and Air Conditioning) equipment or central heating can be controlled according to information collected by various sensors that are monitoring parameters such as temperature, humidity, light or the presence of people in the room division. The use of sensors, such as smoke detectors, motion or breaking glass, allow to detect possible risk situations and initiate appropriate actions. Automated home entertainment systems allow users to predefine favorite settings, remove the need for multiple remote controls and listen to music, news or podcasts in any room of the house.

The home automation systems have a great application in the elderly population range in terms of comfort, convenience and safety. The increase in life expectancy has brought enormous challenges to many aspects of human life, especially in the welfare and health care of the elderly. Some problems associated with reduced mobility and dementia can be minimized through the use of remote control of a home automation system while the use of sensors can alert potential hazards.

Often, the concept of smart home is confused with home automation. The concept of smart home is usually used to define a housing which integrates technology and services in order to increase energy efficiency and improve the quality of life [3].

In this work, it is presented a low cost domotic system based on open hardware and software with support for a wide range of equipment. The system is based on a mini computer Raspberry Pi network that can scale according to the number of devices to be controlled. The system enables the control of these devices with a mobile device, and also allows the definition of `if-then` rules to automate certain types of tasks. It is even possible to configure alerts that warn users when certain events occur. The system also include a consumption register module that provides analysis mechanisms to allow a more rational use of energy.

In addition to this introductory section, this paper is organized as follows: Section 2 presents a summary of key technologies and projects in the field of home automation. Section 3 presents the proposed solution for a home automation system and Section 4 deals with the evaluation performed. Finally in Section 5 we present the conclusions and future work.

II. RELATED WORK

Communication protocols are the language that allows the sensors and actuators of a home automation system communicate with each other. Of all, the X10 protocol is the most popular. It was created in the mid-70s, and uses the electric network to send data between devices, which makes the installation of these devices easier without the need to

 ISBN: 978-1-61208-440-4

pass any cables. There are a number of compatible devices with this protocol and its main drawback is the lack of robustness, because it is unidirectional and does not include a verification mechanism of sent commands. Nevertheless, its popularity led the adoption to work through radio frequency (X10 RF), and it is actually, the most widely used protocol in the world of home automation with numerous devices that support this protocol.

In recent years, there has been a proliferation of the use of wireless networks to control devices, which has facilitated the installation without much intervention on existing infrastructure. These systems, called Wireless Home Automation Networks (WHANs), include built-in sensors and wireless actuators that enable monitoring and control applications for home user comfort and efficient home management [2]. The main solutions of WHANs are based on X10 RF, ZigBee, Z-Wave, INSTEON and Wavenis networks. Those protocols differ in terms of the frequency used, range (indoors and outdoors), speed and costs. Gomez and Paradells [2], presents a survey of WHANs architectures and technologies detailing the main features of those protocols.

One of the most popular open source projects for general home automation is MisterHouse that is customizable, very flexible and fully compatible with most existing technologies [4]. A major problem of this system is it require good Perl programming skills to set it up and run it.

The Node-RED [5] project of IBM, allows to connect electrical devices, APIs and services online using a visual editor to build all logic and system behavior. Links are programmed as a flow chart in a very simple way, according to a targeted approach to the data flow. In fact, the Node-RED is more than a home automation system being considered by IBM a visual tool for wiring the Internet of Things (IoT).

`pilight` [6] is an open source domotic solution that runs on mini computers like the Raspberry Pi, the HummingBoard, the BananaPi, but also in various Linux distributions such as Arch, Ubuntu or Debian. The great advantage of `pilight` is the support for a wide variety of WHAN devices from different manufacturers, provided that they are compatible with the X10 protocol over RF (either 315 MHz or 433 MHz). `pilight` has a web interface that allows the control of the equipment through any browser and there are also several native applications for Android and iOS operating system. Since version 6 that `pilight` provides an `if-then` rule engine that allows the automation of tasks. All the configuration of `pilight` is made through a set of JSON files, which includes among others, the available devices, home divisions and rules.

Baraka et al. [7] present a solution similar to the solution presented in this paper but based on an Arduino and with fewer features or expansion possibilities.

In this work, we use open hardware and software with the module that allows the configuration and management of the system developed from scratch. We use `pilight` described above, and tested the system with the objectives of assessing system performance, test the system features in real situations and evaluate the system architecture.

III. PROPOSED SOLUTION

Fernandez and Losada [8] present a set of requirements quite common and discussed in the literature that aim to guide the design of home automation solutions. These requirements are summarized next, which were considered in the architecture solution proposed here and to which was added the low cost requirement.

1. **Low Cost.** The solution should have a low cost using inexpensive and modular equipment in order to grow with the user needs.
2. **Interoperability**. The central control system must be open and able to communicate with all sensors and actuators, in order to connect to various types of devices. Although some standards were adopted, there must be especially careful when purchasing new devices to ensure compatibility with the existing home automation system.
3. **Control and Remote Access**. The ability to control and communicate remotely with the system, is one of the most popular features. This feature allows users to monitor the home environment and also change the settings of lights, thermostats and other equipment, from a laptop, personal computer, smartphone or any mobile device.
4. **Scalability**. The system chance to expand its capacity according the needs and the evolution of technology itself.
5. **Flexibility**. The system must provide tools to customize the system to specific requirements.
6. **Robustness**. The system must be able to handle the most common problems, such as power failures, data persistence and adequate reporting mechanisms.
7. **Energy Saving**. The system should have the ability to help save energy by automatically turning off electrical devices, as well as providing the user with information that assists in decision-making for a more rational use of energy.

A. Solution Architecture

The proposed solution is based on the architecture illustrated in Figure 1, which consists of the modules described next: the Central Module, the Management Module, the Control Module, the X10 RF Devices and also the X10 RF Power Consumption Measuring Devices. The system enables the control of home automation devices compatible with the X10 RF protocol, thereby providing support for a large number of multi-vendor equipment.

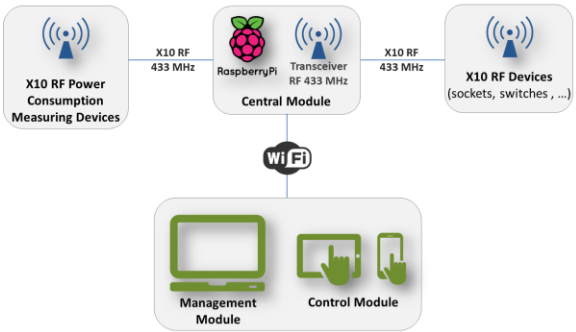


Figure 1. System architecture.

1) *Central Module*

The Central Module is responsible for maintaining the configuration and control of the home automation system. It is the component which communicates with home automation devices sending and receiving commands that allow a rational management of the system, such as turn on/off or adjust the lighting, opening or closing the shutters, turn on the washing machine at a desired time, among others. This module comprises a Raspberry Pi with a 433 MHz RF transceiver for a maximum range of 150 meters. The Raspberry Pi is a low-cost computer (from 20 €) with a size of a deck of cards. It runs Linux (Debian, Fedora or Arch) and has capabilities to interact with the outside world via programmable General Purpose Input / Output (GPIO) ports. The Raspberry Pi runs the `pilight` mentioned above and can be configured with the Management Module described next.

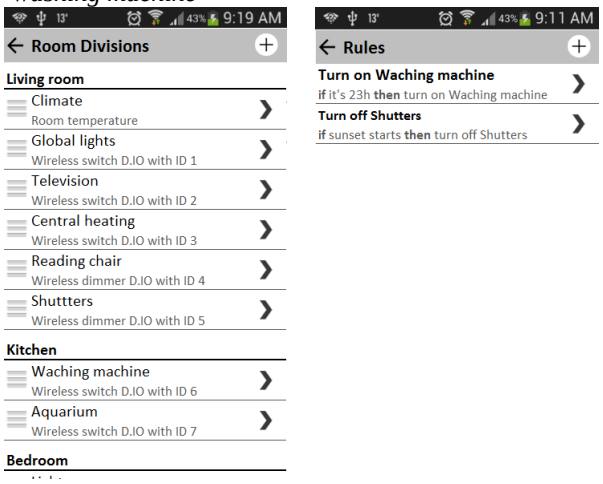
2) *Management Module*

The management module is a web application developed with jQuery Mobile that allows the management and the configuration of the entire system. `pilight` does not have this module, which was developed as part of the system here presented. This module is a web editor for the configuration files of `pilight` which maintains all the settings in JSON files. Figure 2 shows an example with the interface of the configured room devices with the corresponding devices (Figure 2a) and the created rules (Figure 2b).

A `pilight` rule is composed by a condition and an action. The condition defines the triggers that allow the execution of the action associated with the rule when the condition is verified.

The `pilight` rules module is very expressive and the management module allow the creation of conditions, like:

"*If it is 23h then turn on the washing machine*"
"*If it is time for sunset then close the shutters*"
"If the room temperature is greater than 25 degrees then turn on air conditioning"
"If it is 23h of Tuesday or Friday then turn on the washing machine"

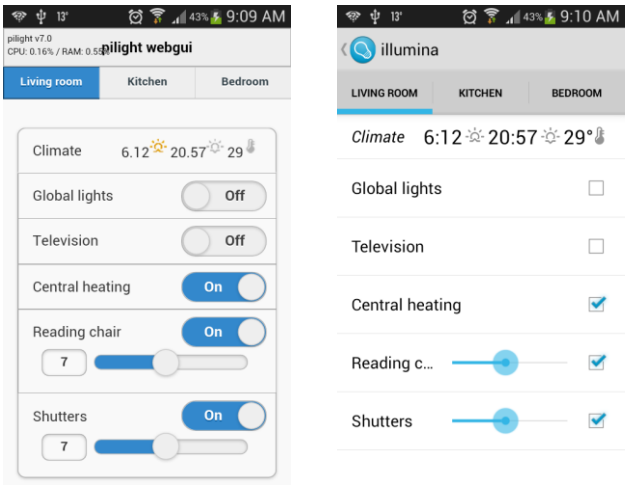


a) Room divisions b) Rules

Figure 2. Management module interface.

3) *Control Module*

`pilight` provides a web interface (Figure 3a) for the control through a mobile device. In addition to the web interface there are also available several native applications for Android (Figure 3b) and iOS that allow the same control. Those interfaces organizes information by the house division that are defined in the Management Module and within each division includes the devices to control. In addition to the domotic devices, it is possible to include information about the weather forecast gathered from available APIs like Weather Underground or Open Weather Map.



a) Web interface b) Android app (illumina)

Figure 3. Control module.

4) *X10 RF devices*

The solution is based on the use of X10 RF devices. We have used D.IO Chacon devices whose unit price starts at 7€. The D.IO product catalog includes sockets, switches, on/off modules, flood detectors, temperature sensors, humidity and motion, among many others.

5) *X10 RF Power Consumption Measuring Devices*

This is a functional module separated from the previous one, but which also use the X10 RF protocol to communicate with the Central Module. The use of power consumption measuring devices provides information that can be used for a more rational use of energy. This information includes current and historic consumption per equipment or room.

IV. EVALUATION

The system was tested in a house with about 150m^2 and domotic devices installed in three divisions of the house: the living room, the kitchen and one bedroom. Was used a total of 12 actuators from DI.O Chacon (9 switches and 3 dimmers) and 1 sensor (the AM2302/DHT22) which read the temperature and humidity of the living room.

For the Central Module was used one Raspberry Pi version B which has a 700 MHz ARM processor and 512MB of RAM. The DHT22 sensor and the RF 433 MHz transceiver was mounted on a breadboard according with the instructions available at the support site of `pilight`.

 ISBN: 978-1-61208-440-4

The Management Module was used to configure all the system including the divisions of the house and the installed devices (Figure 2a), and also the rules used to perform automatically some predefined tasks (Figure 2b).

For the Control Module, was mainly used the Android app illumina (Figure 3b) and less frequently the `pilight` web interface (Figure 3a).

The tests were conducted for a month by the inhabitants of the house (one adult with 48 years old and two young people with 13 and 16 years old) on a regular daily basis with the objectives of (i) assessing the capacity and system performance, (ii) test the system features in real situations, and (iii) to evaluate the system architecture.

A. Results

With the test conditions described above the CPU usage is about 21-33% which is quite acceptable. This use corresponds mainly to decoding the RF signals and filtering and also the `pilight` rules engine that makes more use of CPU. As a result, system load largely depends on the rules and also the sensors used which broadcast RF signals from time to time forcing the use of CPU to decode these signals.

The distributed and scalable architecture of `pilight` allows easily add new Central Modules operating in a master-slave philosophy thus distributing the load of CPU and obtaining a higher range of the system than the 150 meters of the RF transceivers used.

During the tests, the system proved its reliability running without the need for intervention.

V. Conclusion and Future Work

This paper presents a low cost system based on open hardware and software that enables the remote control of devices compatible with the X10 RF protocol. The system can also automate tasks through an `if-then` rule engine and allow the measuring of power consumption.

Each Central Module in prototype version as the tested here, costs about 50 €.

The use of `pilight` as middleware platform for home automation devices, allow to address the interoperability for a wide range of equipment from various manufacturers.

The performed tests allowed to validate the system architecture and has proven to function well in a house of about 150 m^2. However, it is desirable to perform larger scale tests especially with the use of more home automation devices from various manufacturers and an enlarged Raspberry Pi network.

An important issue that will be considered in future, is the security of the system and also tests with the power consumption module.

References

[1] A. J. B. Brush, B. Lee, R. Mahajan, S. Agarwal, S. Saroiu, and C. Dixon, “Home automation in the wild: challenges and opportunities,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11), May 2011, pp. 2115-2124, doi: 10.1145/1978942.1979249.

[2] C. Gomez and J. Paradells, “Wireless home automation networks: a survey of architectures and technologies,” IEEE Communications Magazine, vol. 48, no. 6, pp. 92-101, Jun. 2010, doi: 10.1109/MCOM.2010.5473869.

[3] M. A. E. L. Mowad, A. Fathy, and A. Hafez, “Smart home automated control system using Android application and microcontroller,” International Journal of Scientific & Engineering Research - IJSER, vol. 5, no. 5, pp. 935-939,May 2014, ISSN: 2229-5518.

[4] M. Ali, J. H. A. Vlaskamp, N. N. Eddin, B. Falconer, and C. Oram, “Technical development and socioeconomic implications of the Raspberry Pi as a learning tool in developing countries,” in 2013 5th Computer Science and Electronic Engineering Conference (CEEC), IEEE Press, Sep. 2013, pp. 103-108, doi:10.1109/CEEC.2013.6659454.

[5] IBM. *Node-RED: A visual tool for wiring the Internet of Things*. [Online]. Available from: http://nodered.org/.

[6] *pilight - Open Source Domotic Solution*. [Online]. Available from: http://www.pilight.org/

[7] K. Baraka, M. Ghobril, S. Malek, R. Kanj, and A. Kayssi, “Low cost Arduino/Android-based energy-efficient home automation system with smart task scheduling,” in Proceedings of the Fifth International Conference on Computational Intelligence, Communication Systems and Networks (CICSYN '13), IEEE Press, Jun. 2013, pp. 296-301, doi:10.1109/CICSYN.2013.47.

[8] J. L. Fernandez and D. P. Losada, “An integrated and low cost home automation system with flexible task scheduling,” in 15th Workshop on Physical Agents, Jun. 2014, pp. 70-79, ISBN: 978-84-9773-681-7.

 ISBN: 978-1-61208-440-4

e-Health Control and Location Services for Wandering Patients through Cloud-based Analysis

Javier Rubio-Loyola[1], Gustavo Ramírez[2]
[1]CINVESTAV Tamaulipas, [2]Tairda Innovations
Mexico
{jrubio@tamps.cinvestav.mx,
gustavo.ramirez@tairda.mx}

Tonny Velin[3], Jorge Sánchez[4], Eduardo Burgoa[5]
[3]Answare Technologies, [4]Acesyd, [5]AIDO
Spain
{tvelin@answare-tech.com, jsanchez@acesyd.com,
eburgoa@AIDO.es}

***Abstract*— This paper presents work in progress towards a platform system aimed at providing support in e-health control and provision of location services for wandering patients through real-time medical and mobility information analysis. These critical aspects are pivotal for the PAVISALE project whose core ideas are presented in this paper. The PAVISALE platform provides services for patients, caregivers and for e-health service operators. Patients are enabled to receive notifications and suggestions in response to emergency, contingency situations or deviations from health and mobility patterns. Caregivers receive such notifications as well for further assistance and they are also enabled to analyse health information and mobility behaviour of patients during time scales and territorial scopes. E-Health service operators are enabled to program e-health and mobility control guidelines according to specific needs of wandering patients. A platform system for e-health control and mobility for wandering patients is possible if and only if efficient and robust wearable devices are designed and they are integrated into a holistic analysis system that can monitor, analyse and store medical information in run-time aligned to the needs of wandering patients' health control and mobility constraints.**

Keywords-wandering patients`; Location-based services; medical information processing; cloud.

I. INTRODUCTION

Dementia refers to the loss of cognitive functioning. People suffering from dementia are affected in their ability to think, remember and reason. There are a number of conditions that cause symptoms of dementia. The most common causes are Alzheimer's disease, vascular dementia or multi-infarct dementia with Lewy bodies and frontotemporal dementia [1]. Although vascular dementia covers 20% to 30% of all cases of dementia, Alzheimer's disease is the most common cause of dementia covering 50% to 75% of cases. It destroys brain cells and nerves, affecting the transmitters that send messages to the brain, particularly those responsible for storing memories.

As mentioned by León-Ortíz et al. [2], the age is the most important risk factor for developing a dementia syndrome. It is expected that by 2030, 60 countries will have more than 2 million elderly persons over 65 years. This will have great impact on public health in several countries because dementia is the leading cause of disability in elderly people, accounting for 11.9% of years lived with disability from non-communicable diseases. Symptoms of Alzheimer's disease are the sudden loss of memory, confusion in everyday activities, adoption of disturbing behaviour like getting up in the middle of the night or wander errantly and getting lost [1].

People with dementia are at constant risk of wandering away, being exposed to suffer physical, emotional damage and even death. In addition, wandering patients, mostly elderly people, frequently have other diseases that could demand additional health control processes (e.g., taking pills, blood pressure and temperature checkups, etc.), which in turn could be jeopardized due to their wandering behavior.

This paper presents our work in progress towards a cloud-based platform that supports e-health and location services for wandering patients. We present our initial ideas towards a system that would allow monitor specific biomedical parameters, produce alerts in case of abnormal biomedical behavior or unknown location of wandering patients, and also, to produce suggestions from the system to control the health of the patients, all in all, to increase the quality of life of wandering patients. The novelty of this platform relies on the improvement of the capabilities of e-wearable devices, and its integration with a novel intelligent system, which both together can be exploited to produce comprehensive tracking, monitoring and control of healthy life in wandering patients with specific needs. The system will exploit cloud services to enable intelligent storage and processing of information in real time, in favour of scalable e-health services. The system will allow monitor, locate, detect patterns of behaviour and mobility, and follow up of medical treatments using smart technology for a large number of patients. The proposed system is the target of the PAVISALE project, which is a research and innovation project targeting a platform for assistance of healthy life control and location of wandering patients.

After this introduction, Section II presents the state of the art. Section III presents the conceptual framework of the platform system. Section IV presents the technological impact of the proposed solution, and finally Section V concludes the paper.

II. STATE OF THE ART

The last decade has seen a tremendous interest for all aspects of e-health. As a comprehensive survey would require more than a single paper, this section describes selected work related to the system subject of our study.

Intercom Sick-nurse Systems [3] provides support to have the information and patient's alarms in one centralized system.

 ISBN: 978-1-61208-440-4

It includes breathing equipment and electronic whiteboards where the staff of doctors and nurses has relevant information of the patient, reports the presence of nurses or doctors in the room, and continuously monitors the system for fault detection and alarms to alert staff by means of pagers and wireless phones. The operator can make group calls, monitor and setup service requirements, locate staff members, verify service needs and identify sites with staff availability.

The device I'm Home [4] allows taking vital signs of patients at home through sensors, which are sent to the cloud and then it is reviewed by the physician through mobile devices. When a parameter goes outside the normal range, for example, if the pressure rises above 180 or if the temperature rises above 38 degrees, the alarm triggers a call to the call center, which in turn communicates with the patient through a videoconference. I'm home records temperature, heart rate, blood pressure and oxygen levels. The device can be used by patients recovering from the hospital, people with severe lung disease, cancer patients after chemotherapy and elderly people.

E-doctor [5] is a platform for home monitoring in real-time using bluetooth. The application works with physiological sensors and it uses smartphones as gateway. It allows monitor several patients at once. The application presented in [6] uses a WSN (Wireless Sensor Network) tele-monitoring system for health care. It proposes a milti-protocol architecture with physiological sensors and biokinetic sensors. Is uses multiple gateways and it can handle several patients. The application presented in reference [7] is an experimental application that uses a smartphone as Gateway, the bluetooth protocol, physiological sensors, specifically to monitor heart rate, and Internet to transmit data to a remote center. The application presented in [8] is a pilot application that works with Bluetooth and Zigbee (802.15.4) protocol; it offers a redundant system in case of failure of a protocol. It uses sensors and physiological biokinetic sensors.

The e-Health system presented by Cervantes et al. [9] is a research application based on wireless sensor networks that works with physiological sensors, ZigBee (802.15.4) with bridge to WiFi and a personal computer as gateway. The application Cardiosentinal presented by Gao et al. [10] is a sensors application for remote monitoring of cardiac signals with smartphones that is active 24 hours a day. It allows monitoring several persons with cardiac problems and it is run on an infrastructure that allows to act at a given moment in case of a medical emergency. The application presented by Hu et al. [11] is one of the classic applications along with others like "Codeblue" and MEDiSN, which were the precursors of medical treatments and disease control technology, specifically with the use of sensors. This application has evolved to real time streaming using different protocols and different sensors.

A detailed study of sensor networks and applications oriented to collect and analyze biomedical signals is presented by Campaña-Bastidas and Londoño-Peláez [12]. The characteristics of the applications described above (and others in the literature) have a common denominator; the information collected through sensors takes place in real time. However, only a few consider the operation and evolution of the monitored parameters to predict emergencies, or they do not perform data analysis to generate alerts, recommendations or suggest any specific treatment.

From the information included in this section and other works in the literature, we acknowledge that there is a tendency for the use of smartphones to capture and transmit data from sensors. However, the most important limitation of most of the works related to our work is that the use of real-time information analysis is not yet predominant for giving suggestions for health control. In addition, none of the applications consider monitoring the location together with health control approaches, both optimized with the aim at enhancing the quality of like for wandering patients. These basic aspects are of particular relevance to the PAVISALE project.

III. CONCEPTUAL FRAMEWORK

The general objective of our work is the design, development and implementation of a permanent control system for patients that need special care, namely patients that suffer from disorientation disorders or diseases such as senile dementia, schizophrenia or Alzheimer.

The Figure 1 shows a graphical representation of the system platform that we are targeting in our research.

The proposed solution includes the following elements:

• Users' (patient) wearable devices that allow capturing physical activity signals and special biomedical parameters for wandering persons.

• Wireless access technology to support ubiquitous transmission of such information to the cloud.

• Localization infrastructure based on GPS via Wi-Fi or other wireless network.

• Storage and real-time analysis services in the cloud. The cloud will be used to store raw biological and location information of the patients. Data mining and learning algorithms will be executing in the cloud to track, analyze and to determine whether the patient is within appropriate biological levels and also whether the patient is physically located in areas that could fall within his/her location patterns. Algorithms development in this phase are critical because the platform will deal with a large number of patients and also, with specific needs. Scalability passes through efficient and scalable analysis mechanisms that will be addressed in this phase of our developments.

• Real-time self-learning system with manifold objectives: i) to determine patients' behaviour parameters; ii) to monitor for thresholds crossing of biomedical parameters; iii) to give suggestions for healthy life aligned to patients' medical state and evolution.

• Web and mobile applications for users; this is caregivers and/or residence personnel to follow-up, monitor, and react to incidents or events related to the health of the patients.

• A web service interface where administrators can program, and update health control parameters and follow-up guidelines for each patient.

The end-users (patients) will be provided with a wearable device that will implement diverse sensors for vital signal measurement. The capabilities of the wearable device will be aligned to the needs of the patients. The information from the devices will be transmitted to a cloud storage service through a wireless access point. The system will be designed to be aware of the capabilities and nature of the wearable devices and it will

 ISBN: 978-1-61208-440-4

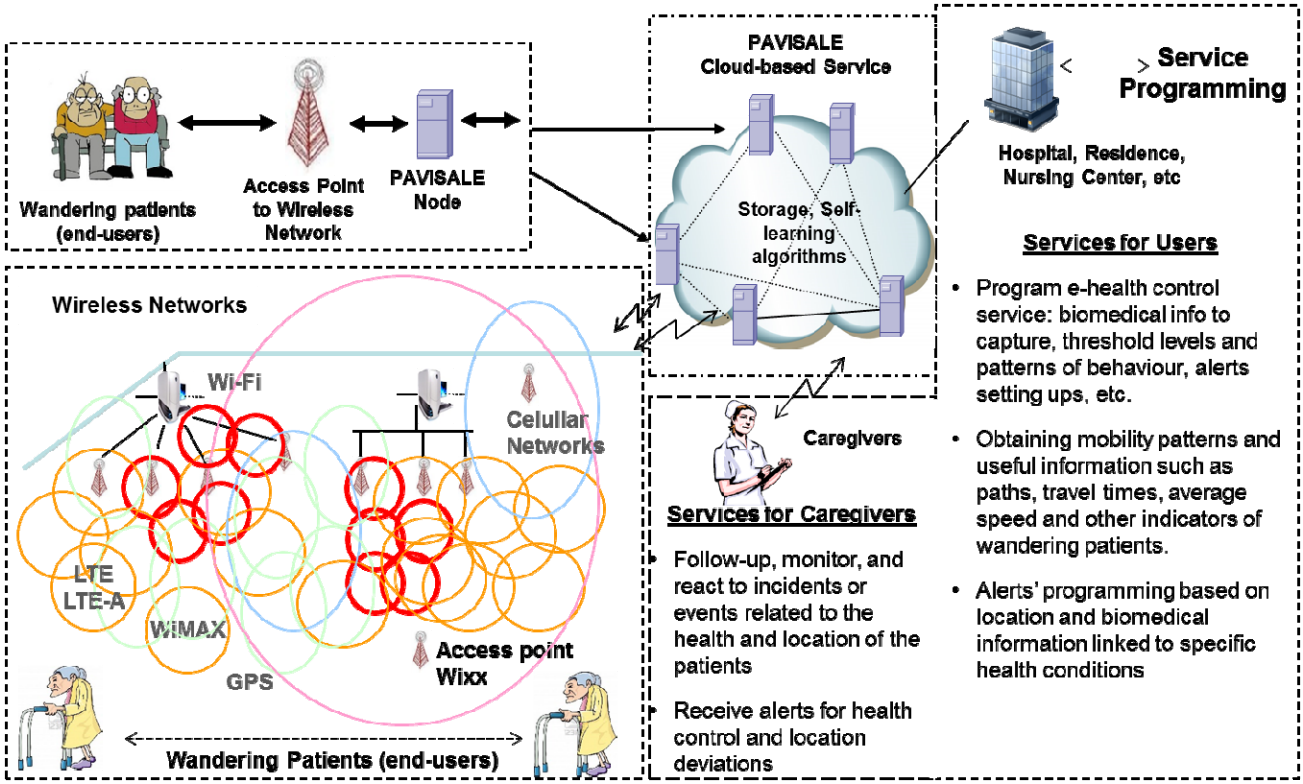


Figure 1. Services supported by the location-based service platform

implement appropriate gateway transmission devices coupled with the sensors.

The users of the platform (caregivers, residence/hospital staff) will have at hand a localization infrastructure that will follow-up, monitor, and react to incidents or events related to the location of the patients.

The cloud service will gather and store the patients' information making use of appropriate anonymization techniques or encryption security protocols to ensure confidentiality.

The self-learning component will analyse the biomedical information in order to detect anomalies as well as to suggest new activities and reinforce those that have produced good results in the patients. The self-learning component will provide services for analysis, recommendation and alert generation. Alerts will be sent to the specialized staff of the residence (end user) through a web application that could run in tablets, PCs, smartphones, etc. Also, the patient could receive the alert if he/she is capable of reacting to such alert accordingly.

To date, there are two types of products related to the solution described earlier: i) systems for locating wandering persons [13] and ii) wrist bands for physical activity registration [14][15]. The proposed solution covers completely the functionality of systems for locating wandering persons, and it greatly enhances the functionality of the wristbands for physical activity registration. Our solution does not only measure the taken steps, but it also offers real-time measurement of sensitive vital signs for wandering persons which comprise temperature, heart rate, skin humidity or breathing. Also, the main innovation that differentiates our approach from other products is that it incorporates a system that analyses the activity and parameters of all users (patients), offering tips to follow a healthy life style, and the auto prediction of possible health problems. In the next section, we elaborate on the technological challenges and contributions to all actors involved in the solution presented in our work.

IV. TECHNOLOGICAL IMPACT

The expected growth on the quantity of wandering patients for the next years demands the deployment of new solutions that help enhance the quality of life of persons with wandering behaviour, and also that can provide support for more efficient attention and follow up for wandering patients in health centres, residences and hospitals.

This is not an easy task because wandering patients can be elderly, and they may present a combination of various health conditions that require personalized care and continuous monitoring. The constant monitoring of diverse biomedical parameters and the periodic tracking of biomedical information of several wandering patients is only possible by means of the use and development of diverse technologies that include: design of electronic devices with restrictions of usage, portability, and energy efficiency, a safe and plug-and-play communication protocol that guarantees the smart gathering of biomedical data, the development of new scalable storage and information analysis services, and finally, smart tools that allow hospital directors/administrators and technical staff of residences to track physical location, as well as monitor the evolution of biomedical parameters of patients. The contributions to the state of the art in all actors involved in the solution proposed in this paper are described hereafter.

 ISBN: 978-1-61208-440-4

A. Impact on health care operators

The contributions to health care operators are manifold and they are summarized in this section.

The system will extend the health care operators' role to a more active role that is commonly taken by specialized service operators at expensive costs. In this sense, public and private hospitals as well as diverse residencies with wandering patients will be enabled to define and operate autonomously the health care services for their communities according to the parameters, and the desired tracking for each patient.

The system will allow hospitals, residences, etc., acquire and define patterns of e-health services, which could be instantiated in several places for patients with similar needs. This would also have an impact on the reduction of operation and maintenance costs in e-health care services.

Cost reduction has a domino effect on critical aspects of e-health services provision, such as: i) increase the offer of service operators; ii) improve existent services by integrating new devices for the monitoring of biomedical signals with improved capabilities.

The proposed system will be designed to be flexible and scalable allowing for the integration of new specialized or emerging monitoring devices.

The system will be tested in real environments. Moreover, an important effect of its application in real environments is that it will bring new business models for its exploitation in commonly neglected environments such as rural residencies, regional hospitals, and remote homes with patient tracking needs.

B. Impact on device manufacturers

The proposed system will allow manufacturers of electronic/wearable devices to get know-how for developing emerging technologies for monitoring and tracking healthy life of wandering patients. In addition, they will design and develop more and better devices that could be integrated to platforms of monitoring and analysis of large volumes of information. The above will have an impact towards mass production of devices with improved capabilities. Finally, device manufactures will adapt to market evolution to address the requirements of the increasing number of wandering patients in the world.

C. Impact on wandering patients

The proposed system will have a significant impact on wandering patients, since the researched technologies in this work will allow the provision of services for systematic location tracking and biomedical data monitoring in general. Wandering patients could enjoy such advanced services at affordable prices and in consequence an increase in its quality of life due to the assistance and personalized tracking of location and health control.

D. Impact on family members

The family members of wandering patients would benefit from our platform system in several ways. They would be able to monitor and pay special attention about biological information and physical location of their relatives through their smartphones, tables, etc. With this regard, family members could be warned when the patient may experience deviations of biological state or in case the patient gets lost at a given time. With the support of the platform system, family members could call for assistance or they could be able to provide the caregivers access to the patient in cases when the latter ones attend the alerts triggered by the system. Family members could enjoy these services at affordable prices and they would increase their quality of life as they would be supported by an intelligent and reliable system when looking after their wandering relatives.

V. CONCLUDING REMARKS

This paper has presented work in progress towards a platform system aimed at providing support in e-health control and provision of location services for wandering patients through real-time medical and mobility information analysis. The conceptual framework presented in this paper indicates that e-health control and location services for wandering patients can be achieved with appropriate enhancements of wearable devices that can transmit biomedical information towards an analysis system that can monitor, analyse and store medical information in run-time aligned to the needs of wandering patients' health control and mobility targets. This paper has presented the core ideas with which we think that we will contribute to the state of the art in the critical nature of providing support in e-health control and provision of location services for wandering patients. Mainly, we expect to contribute to the state of the art with an implementation of the conceptual design presented in this paper, and the specific algorithms that would analyse and process the large amount of biological information to enhance the quality of life of wandering patients.

ACKNOWLEDGMENT

This work is partially supported by the Spanish Government by its funding through the project COSAIF, "Connectivity as a Service: Access for the Internet of the Future" (TEC2012-38754-C02-01 and TEC2012-38754-C02-02) and the Council for Science and Technology of Mexico (CONACyT) under the CONACYT-CDTI Project 189413.

REFERENCES

[1] Alzheimer's Disease International, "Types of dementia," The Global Voice on Dementia, http://www.alz.co.uk/ visited last time Oct 6, 2015

[2] P. León- Ortíz, M. L. Ruiz-Flores, J. Ramírez-Bermudez, and A.L. Sosa-Ortíz. Lifesyle in elderly people and their association with dementia. Gaceta medica de Mexico 149(1):36-45 · July 2013, pp. 36-45

[3] "Sistema Enfermo-Enfermera." [Online]. Available: http://www.adix.com.mx/Productos/Enfermo-Enfermera.html. [Accessed: Oct 6, 2015]

[4] Article "Making life easier for remote patient monitoring" [Online]. Available: http://www.am.com.mx/leon/tecnologia/facilitan-monitoreo-de-pacientes-a-distancia-114345.html. [Accessed: Oct 6, 2015]

[5] Z. Jin and Z. Qian. E-doctor: A real time home monitoring and mobile healthcare plataform. IEEE ComSoc Multimedia Communications Technical Committee Letter, vol. 6, no. 11, November 2011, pp. 36-38

[6] J. M. Corchado, J. Bajo, D.I. Tapia, and A. Ajith. Using heterogeneous wireless sensor networks in a telemonitoring system for healthcare. IEEE Transactions on Information Technology in Biomedicine. vol. 14, no. 2, March 2010, pp. 234-240.

[7] P. Crilly and V. Muthukkumarasamy, "Using smart phones and body sensors to deliver pervasive mobile personal healthcare," Sixth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), Australia, 2010, pp. 291–296.

 ISBN: 978-1-61208-440-4

[8] F. Felisberto, N. Costa, F. Fdez-Riverola, and A. Pereira, “Unobstructive Body Area Networks (BAN) for Efficient Movement Monitoring,” Sensors, vol. 12, no. 12, Sep. 2012, pp. 12473–12488,

[9] H. Cervantes, J. I. Nieto, J de D. Sánchez, M.E. Martínez, and A. H. Calvo. e-Health architecture based on wireless sensor networks. DIF100ci@, ISSN 2007-3585, vol. 6, no. 2, December 2012. Pp. 54-61

[10] M. Gao, Q. Zhang, L. Ni, Y. Liu, and X. Tang, “CardioSentinal: A 24-hour Heart Care and Monitoring System,” J. Comput. Sci. Eng., vol. 6, no. 1, Mar. 2012, pp. 67–78

[11] F. Hu, M. Jiang, L. Celentano, and Y. Xiao, “Robust medical ad hoc sensor networks (MASN) with wavelet-based ECG data mining,” Ad Hoc Networks, vol. 6, no. 7, Sep. 2008, pp. 986–1012

[12] S. E. Campaña Bastidas and J. M. Londoño Peláez. Wireless Sensor Networks and Related Applications for Collecting and Analyzing Biomedical Signals. Gerenc. Tecnol. Inform. vol. 12, no. 33, May - Ago, pp. 85 – 99

[13] M. Avvenuti, C. Baker, J. Light, D. Tulpan, and A. Vecchio, “Non-intrusive Patient Monitoring of Alzheimer’s Disease Subjects Using Wireless Sensor Networks,” World Congress on Privacy, Security, Trust and the Management of e-Business, 2009, Canada, pp. 161–165

[14] “Medical Alert Systems & Service | Philips Lifeline ®,” Philips Lifeline. [Online]. Available: http://www.lifelinesys.com/content/. [Accessed: Oct 6, 2015]

[15] “Medical Alert Systems | Medical Alert Services for Seniors - Alert1®.” [Online]. Available: http://www.alert-1.com/. [Accessed: Oct 6, 2015]

 ISBN: 978-1-61208-440-4

Components Interconnection Consideration In Multi Metrics Approach

Seraj Fayyad

University of Oslo/UNIK
Oslo, Norway
Email: seraj@unik.no

Josef Noll

University of Oslo/UNIK
Oslo, Norway
Email: josef@unik.no

***Abstract*—Internet of Things, People and Services (IoTPS) systems have become increasingly popular in modern times. And this popularity increases the importance of measurable Security, Privacy, and Dependability (SPD). One of the crucial aspects for system SPD enhancement is reliable evaluation for system SPD level. The evaluation of SPD level for IoTPS system has many challenges, such as the heterogeneity among the components. Considering the challenges of IoTPS system, several approaches are proposed to evaluate system SPD level. One of these approaches, is Multi Metrics (MM) approach. This approach is considered as comprehensive approach, because of its features. Some of MM approach features target the scalability and applicability within the architecture of unlike systems. To enhance the comprehensiveness of MM approach, we propose an extension for the approach to consider the impact of components interconnection on SPD level.**

***Index Terms*—components interconnection; interconnection weighting equation; mHealth system; IoTPS; Multi-Metrics; security level; privacy level; dependability level.**

I. INTRODUCTION

Internet is transforming from communication highway between computers into a backend system connecting hybrid networks. Within these hybrid networks, people, services, things (sensors, actuators) and computers are connected as one. A good example for IoTPS systems is smart grid system, which consists of diversity subsystems. The interaction of Subsystems provides powerful services, such as grid monitoring and remote controlling.

Another kind of IoTPS systems is mHealth systems. This type of system defined by Adibi as the practice of eHealth assisted by smartphones, which are used to capture, analyze, process, and transmit health-based information from sensors and other biomedical systems [1]. Some of provided services by this kind of systems are regular monitoring, real time advising, auto-notification in emergency cases and also affection of emotional states as stated by Cipresso et al. [2].

Eloff et al. envisaged that an IoTPS system will require focus on security and privacy [3]. Despite of many advantages of IoTPS systems, these systems arise new security, privacy and dependability concerns. One of these concerns, is the heterogeneity among subsystems, which complicates system SPD evaluation and satisfaction. Another concern, is the new open area for exploitation of the system, such as the mobile of the patient in mHealth systems.

Garitano et al. propose a Multi Metrics (MM) approach, being a comprehensive and dynamic approach for the evaluation of SPD level for a given IoTPS system [4]. They demonstrate the MM approach applicability by performing it on the smart vehicle IoTPS system. Noll et al. demonstrate more features of the MM approach, such as applicability on huge IoTPS system (e.g smart grid) and scalability [5].

This paper enhances the MM approach, by considering the impact of interconnection on the SPD level. The paper is organized as follows: In Section II, we give an overview of related work. In Section III, we elaborate the proposed extension for MM approach. In Section IV, we demonstrate proposed extension, by applying it on mHealth system as use case. In Section V, we present our conclusion.

II. RELATED WORK

Different approaches have been developed for analyzing of IT system risks. Based on the envisaged focus, Manadhata and Wing classified these approaches into attacker-centric approaches and design or system-centric approaches [6].

Attacker-centric approaches, are based mainly on the knowledge about the system attacks. Usually, these approaches collect and analyze attacks-related data, such as; system vulnerabilities, goals of system attackers and detected malicious activities. For the collection of attacks-related data different resources could be used, such as; Intrusion Detection System (IDS) and National Vulnerabilities Database (NVD) [7]. Based on collected data, these approaches build system-attacks model, to analyze the risks of IT system.

The most popular attack models are: attack graph [2], [8], attack trees [9] and Bayesian network [10]. Wang et al. propose an attack graph-based probabilistic model, to quantify the security of IT system network [11], [12]. Wang et al. propose attack graph analysis to be used as a knowledge base for correlating IDS received Alerts, hypothesizing missing alerts, and predicting future alerts [13]. Xie et al. use Bayesian networks incorporated with IDS alerts to analyze the security risks of the IT system [10]. Schneier proposes the analysis

 ISBN: 978-1-61208-440-4

of system risks using knowledge about attackers coupled with attack trees [14]. Dantu et al. propose the usage of attack graph coupled with the behavior of attacker to analyze the risks of IT system [15].

System-centric approaches concentrate on system design and architecture for risk analysis. Manadhata and Wing propose methodology for software system's attack surface measurement [6]. The concept of proposed methodology considers attack surface comparable to system security level (the smaller the surface, the more secure the system)

Howard et al. propose attackability metric to measure system security level, through the measurement of system attack surface from three dimensions [16]. These dimensions are: targets and enablers, channels and protocols, and access rights. Howard et al. refer to the increasing of attack surface and reduction of security level, caused by increasing of targets, channels, and generosity of access rights.

Garitano et al. propose MM approach for the evaluation of SPD level for a given IoTPS system [4]. One of the important MM approach features is the comprehensiveness. It starts with component evaluation, then sub-systems evaluation and ends up with the entire system evaluation. Garitano et al. demonstrate that, different configurations cause different SPD level. Which demonstrate the possibility for SPD level enhancement, through changes in system configurations. Noll et al. demonstrate the scalality and applicability of MM approach, by applying it on large and complex system, such as smart grid [5].

Interconnection has a general impact on system SPD and component SPD, in particular. For instance, a successful attack on the monitoring compontent of a smart vehicle will exploit the privacy of the vehicle rider. Fayyad and Noll state some examples, which reflect the impact of interconnection on the SPD level [17]. The evaluation of the MM approach was not sensitive to interconnection, thus the failure of components with high interconnection is not appropriately considered. In this paper, we introduce an extension for MM approach, which addresses interconnection impacts. We demonstrate proposed extension by applying it on mHealth system. Which shows, that similar SPD level for a given components could vary based on the interconnection of these components.

III. Multi Metrics APPROACH extension

A. Multi Metrics Approach

The MM approach is system-centric approach for the evaluation of SPD level for a given system. As Garitano et al. [4] and Noll et al. [5] have given comprehensive overviews on the MM approach, this paper concentrates on the effect of interconnections of components.

A security evaluation using the MM approach assumes a hierachical architecture for the system of systems. As shown in Figure 1; the evaluation starts at component level, then evaluates the subsystem and finally addresses the entire system, resulting in a system SPD level.

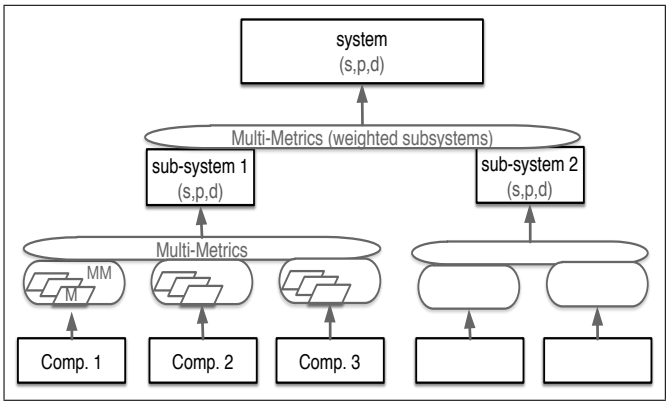


Fig. 1. MM evaluation hierarchical levels, with M indicating a Metrics analysis

In the MM evaluation, the SPD level for an evaluated entity is represented as a triple (S, P, D). Each element of the triple should be described by a value range from 0 to 100. Although, the evaluation ends up with SPD level, SPD criticality is used during the whole evaluation process. From a technological point of view, criticality of a system or a component is more easy to address. SPD criticalities or (Criticality of Security, Criticality of Privacy, and Criticality of Dependibility) (Cs, Cp, Cd) reflect the operation condition like ideal, good, acceptable, or failure. SPD criticalities are defined as complement to SPD triple,

$$(Cs, Cp, Cd) = (100, 100, 100) - (S, P, D). \quad (1)$$

At the components level, the SPD level for a given component is evaluated using a set of metrics. The identification of metrics is performed based on the expected impact of components on a given service. At a later stage, we foresee a framework of metrics being used for evaluations following the MM approach. The final goal is an integrated evaluation for a system, consisting of sub-systems and components.

The MM approach combines the sub-systems using Root Mean Square Weighted Data (RMSWD)(2). The RMSWD formula consists of two parameters, a weight parameter w and a criticality parameter x. Parameters values are within the (0-100) range. In the MM approach, the weight parameter represents the significance of the component or a subsystem on the behaviour of the total system. The criticality parameter represents the operation condition of a component of a subsystem, being e.g. ideal, good, critical or failure.

$$C = \sqrt{\sum_i \left(\frac{x_i^2 w_i}{\sum_i^n w_i} \right)} \quad (2)$$

To summarize the MM approach; the evaluations of metrics for a component are integrated to find the criticality of a component, expressed in terms of security, privacy and dependability. Later, the SPD criticalities of components for a subsystem are integrated using RMSWD and result in a subsystem SPD criticality. Lastly, SPD criticalities of subsystems are integrated

 ISBN: 978-1-61208-440-4

using RMSWD to produce the criticality of a system. The (S, P, D) of a system is then calculated using equation (1).

B. Interconnection Consideration and Positioning

Component interconnection causes reduced security, privacy or dependability, as the failure of one component will not only affects the respective sub-systems, but all sub-systems connected to the component. Let us consider the case of three components (A, B, C) and their interactions. For instance, a body sensor C, which send data to mobile application A, over a Bluetooth connection B. Successful attack on A authentication could enable attacker to reveal the transmitted data confidentiality over B, although it is fully confidential during transmission. Also, a successful attack on A authentication, could enable attackers to inject some malicious scripts, impacting the transmission protocol, which make B or C unavailable. Thus, the SPD of component A impacts the SPD of components B and C and vise versa.

To consider the impact of interconnection on SPD levels, we propose a default metric for interconnection evaluation. The proposed metric is positioned on a level between component and subsystem in MM architecture. Figure 2 shows the proposed positioning of interconnection metric as part of the MM approach.

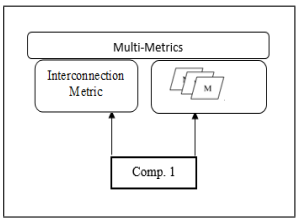


Fig. 2. Interconnection metric positioning within MM

The reason of positioning the interconnection metric in a higher level is, that interconnections impact all components and their SPD parameters. Thus, the evaluation of interconnection for a given component is integrated with component SPD level, where, Formula (2) is used as integration operator.

C. Interconnection Metric

Fayyad and Noll consider system, as a set of interconnected components interact through message passing and/or controlling [17]. Based on this consideration, Fayyad and Noll define an interconnection graph for system components (Definition 1), which model the interconnection and interaction of system components.

Definition 1: Given a set of Components C, having a set of control relations, $Rc \subseteq C \times C$, and a set of data relation $Rd \subseteq C \times C$, then the components interconnection graph G is the directed graph $G(C, Rc \cup Rd)$ (C is the vertices set and $Rc \cup Rd$ the edge set).

Directed graph meant, that the relations or edges between components have a direction associated with them. This direction could be in two way or one way, such as relation between control unit and actuator, which has one way direction from control unit to actuator.

By analyzing of interconnection graph for system components, Fayyad and Noll propose interconnection-based weighting algorithm [17]. The proposed algorithm ends up with a weighting equation (3), which weights the interconnection for specific component.

$$W(c) = S + R + 2C_T + D + A + (\sum_{v=0} \frac{1}{DIST(c,v)}) + V_R \quad (3)$$

- c: targeted component for weighting, where $c \in C$ and C, is set of all components in the system.
- R: number of sub/system components reachable from weighted component c.
- S: number of components reachable from weighted component c through one edge within system interconnection graph. In other words, surrounded component for component c is a component, which interact directly with c without any intermediate component.
- V_R: number of system valuable or key components, reachable from c.
- C_T: number of control relation or edges between weighted component c and other system components.
- D: number of data relation between weighted component c and other system components.
- A: component c activation rate.
- v: valuable component within the system.
- $DIST$: number of components between c and v components.

Activation rate is considered as the frequency of component activation per time unit. Where, activation rate value ranges from 0 to 1. The higher the value the more active the component. Thus, 0 is when, the component is totally inactive, and 1 when, the component is continuously active. When the component is not active, the SPD level of the component does not impact the SPD levels of interconnected components. On other hand, if component is continuously active, then, its SPD level is continuously impacts the SPD level of interconnected components. This mean, activation rate for a given component influences the impact of other parameters in equation (4) on the SPD level of interconnected components. Thus, we propose new optimization for equation (3) for the use in proposed extension. The optimization is represented in equation (4), in which, activation rate is considered as multiplication factor, instead of summation factors. Interconnection metric based on equation (4) to measure interconnection weight for a given component.

$$W(c) = A(S + R + 2C_T + D + (\sum_{v=0} \frac{1}{DIST(c,v)}) + V_R) \quad (4)$$

Based on the need, for equal evaluation of components activation rate, evaluation rate for all components should be evaluated to the same time unit. Thus, a system engineer should evaluate activation rate of component to the smallest time unit. For instance, Let us consider the case of a system

 ISBN: 978-1-61208-440-4

with two components $c1$ and $c2$. $c1$ activation rate is $60/h$ and $c2$ activation rate is $5/min$. For equal interconnection evaluation, activation rate for $c1$ should be used as $1/min$ and for $c2$ as $5/min$.

D. Interconnection Metric Calibration

In MM approach, the SPD level for a given component is represented with a value range from 0 to 100. On the other hand, interconnection metric based on equation (4) resulted value, could vary based on system architect out of (0 to 100) range. Thus, for the integration of the result from interconnection metric with component SPD level, equation (4) result should be calibrated to (0 to 100) range.

To calibrate equation (4) resulted value to (0-100) range, we define architecture-based reference value, as representing maximum weight of interconnection for a component within given system. Let us consider m as a component within given system, and its interconnection weight is the maximum weight. Then, the values of its parameters in equation (4) will be as follow:

- A = 1.
- R = number of all system components(max number of components reachable from m).
- S = number of all system components(max number of components surrounded m).
- C_T = number of all control edges within system interconnection graph.
- D = number of all data edges within system interconnection graph.
- V_R = number of all valuable components within the system.

By having maximum weight of interconnection, the weight of component 'c' could be calibrated to (0-100) range, as shown in equation (5).

$$Weight_c(c) = \frac{Weight(c)}{Weight(m)} \quad (5)$$

Where:
c: given component within the system.
m: most interconnected component within the system.
$Weight_c(c)$: calibrated weight for a value within (0 - 100).
$Weight(c)$: component evaluated weight using equation (4).
$Weight(m)$: system maximum weight of interconnection.

IV. USE CASE

This section demonstrates the proposed extension for MM approach, using the evaluation of Privacy level SPD_p as one of the SPD triple for two selected components. At the start, it gives an overview about Gravid+ system. subsequently, it performs high level analysis for Gravid+ system based on MM approach with concentration on two components. Later, it measures the interconnection of the two components using interconnection metric. Lastly, it evaluates SPD_p level for the two components under different configurations based on MM approach and proposed extension.

A. Gravid+ System:

Gravid+ system elaborated by Garnweidner et al. aims to monitor blood sugar levels in pregnant women with gestational diabetes and assists them with follow up care, such as; diet and exercise [18]. The architecture of Gravid+ system shown in Figure 3 consists mainly of glucometer device interacts with mobile device over Bluetooth connection. The mobile device, host mobile app, which processes and saves glucometer submitted measurements.

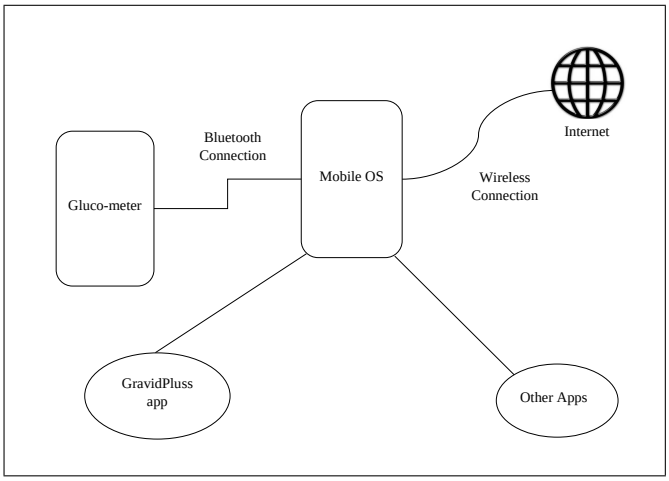


Fig. 3. Gravid+ architecture.

Sub/systems or components, such as other mobile apps hosted on the patient mobile may impact the SPD level of the system, although its not apart of the system. For instance, from privacy perspective, through internet, an attacker could violate one of the mobile apps vulnerabilities and gain access to mobile device. By gaining access to the mobile device, the attacker could violate Gravid+ data confidentiality. Thus, for the analysis of system SPD level, all sub/systems impact the SPD level of the system should be included in analysis.

B. System Components and Metrics

To simplify the demonstration of proposed extension for MM approach, we apply high-level analysis of Gravid+ system. In this analysis, each illustrated nodes within Figure 3 is considered as component. For each of these components a set of metrics are defined as shown in table I. To concentrate on the impact of proposed extension on SPD level, we concentrate on two components (Mobile OS and Gravid+ app), which have different interconnection within Gravid+ system. Where, we assume the performing of similar metrics (authentication and encryption) on the two components,

The analysis of the two components considering SPD_p only. As already stated, MM approach concludes with the SPD level, but SPD criticality is used during the entire evaluation process, as it is shown within tables II to V. The metrics which used in the analysis are:

- Gravid+ app metrics

TABLE I. SYSTEM COMPONENT MAXIMUM POSSIBLE WEIGHT.

Component	*Metrics*
Glucometer	Authentication pairing metric
Bluetooth connection	Encryption metric
Mobile device OS	Encryption metric, Authentication metric
Other apps	number of apps metric
Gravid+ app	Encryption metric, Authentication metric

– Authentication metric(w=40): evaluates the Cp resulted from authentication activation by accessing of Gravid+ app or not (Evaluation is shown in Table II).

TABLE II. GRAVID+ APP AUTHENTICATION METRIC.

Parameters	*Authentication ON*	*Authentication OFF*
Cp	30	70

– Encryption metric (w=30): evaluates the Cp resulted from having App data ciphered or not (evaluation is shown in Table III).

TABLE III. GRAVID+ APP ENCRYPTION METRIC.

Parameters	*Encryption ON*	*Encryption OFF*
Cp	10	60

- Mobile OS metrics
  – Authentication metric(w=40): evaluates the Cp resulted from having authentication activated to access the mobile device or not (evaluation is shown in Table IV).

TABLE IV. MOBILE OS AUTHENTICATION METRIC.

Parameters	*Authentication ON*	*Authentication OFF*
Cp	30	70

– Encryption metric(w=30): evaluates Cp of having data ciphered by Mobile OS or not, using service, such as $encryptdevice$ service in Android OS. One of such service benefits is data protection from offline revealing. (evaluation is shown in Table V)

C. Performing of Interconnection Metrics

To weight the interconnection for a component, using equation (4), the interconnection graph for system components should be initiated. Based on this graph, the values of parameters in equation (4) are driven. Figure 4 shows initiated interconnection graph for Gravid+ system.

Based on the interconnection graph of Gravid+ system, values of interconnection parameters for mobile OS are driven. (shown in Table VI), based on these values, interconnection weight of mobile OS is 23.5.

Parameters of Gravid+ app and their values, are shown in Table VII, based on these values, interconnection weight of Gravid+ app is 12.

TABLE V. MOBILE OS ENCRYPTION METRIC.

Parameters	*Encryption ON*	*Encryption OFF*
Cp	10	60

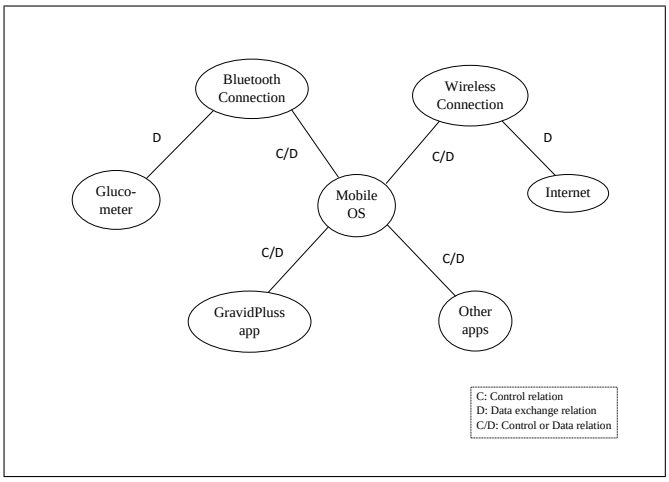


Fig. 4. Gravid+ interconnection graph.

For the calibration of $Weight(c)$ for a given component 'c', system maximum weight of interconnection should be calculated. Table VIII shows parameters of maximum weight and their values in Gravid+ system, based on these values, maximum weight is 28.

By performing equation (5), mobile OS $Weight_c$ is 84 and Gravid+ app $Weight_c$ is 43.

D. Evaluation

Table IX and table X respectively show the evaluation of Gravid+ app Cp and mobile OS Cp under different configurations. $Cp1$ represents the evaluation of configurations based on MM approach. $Cp2$ represents the evaluation of con-

TABLE VI. MOBILE OS INTERCONNECTION PARAMETERS.

Parameters	*Value*	*Note*
S	4	Mobile OS surrounded by 4 components
R	6	Mobile OS has 6 reachable component
V_R	1	Valuable component is only Gravid+ app
C_T	4	Mobile OS has 4 control relations or edges
D	4	Mobile OS receives and sends data from/to 4 components
A	1	Assumption it is continuously active
DIST	2	The distance to the (Gravid + app) is 2 two components (Mobile OS, Gravid+ app)

TABLE VII. GRAVID+ APP INTERCONNECTION PARAMETERS.

Parameters	*Value*	*Note*
S	1	Gravid+ app surrounded by Mobile OS
R	6	Gravid+ app has 6 reachable component
V_R	1	The component is Valuable itself
C_T	1	has one control relation with Mobile OS
D	1	Gravid+ app receives and sends data through Mobile OS
A	1	Assumption it is continuously active
$DIST$	1	Min distance to the app itself

 ISBN: 978-1-61208-440-4

TABLE VIII. SYSTEM COMPONENT MAXIMUM POSSIBLE WEIGHT.

Parameters	*Value*	*Note*
S	6	Supposing that all system components within system surround objective component
R	6	Reachable component in two way from objective component.
V_R	1	Valuable component is only Gravid+ app
C_T	4	Number of control edges within, interconnection graph are 4 .
D	6	Number of data edges within, interconnection graph .
A	1	Component is continuously active
$DIST$	1	Component is valuable by itself

figurations based on MM approach enhanced with proposed extension. Thus, $Cp2$ values resulted from the integration of $Cp1$ with the output of interconnection metric for the component ($Weight_c$). Where, for the two components, the weight of interconnection metric is considered as 30, and the weight of component ($Cp1$) as 70.

TABLE IX. GRAVID+ APP EVALUATION

Encrypt-/Auth-	*M1*	*M2*	*Cp1*	*Cp2*	P_L
Conf.(ON,ON)	10	30	24	31	69
Conf.(ON,OFF)	10	70	53	50	50
Conf.(OFF,ON)	60	30	45	44	66
Conf.(OFF,OFF)	60	70	66	60	40

TABLE X. MOBILE OS EVALUATION

Encrypt-/Auth-	*M1*	*M2*	*Cp1*	*Cp2*	P_L
Conf.(ON,ON)	10	30	24	50	50
Conf.(ON,OFF)	10	70	53	64	36
Conf.(OFF,ON)	60	30	45	59	41
Conf.(OFF,OFF)	60	70	66	72	28

As shown in Table IX and Table X, the $Cp1$ results, of an evaluation of two components lead to the same value. On other hand, $Cp2$ are differentiated based on the interconnections of each components. Thus, the SPD_p will differ. This lead to system SPD level differentiation based on internal interconnection of this system.

V. CONCLUSION

This paper considers systems of systems in the Internet of People, Things and Services (IoPTS). It provides an extension of the Multi Metrics approach including interconnections of components in the system. The Multi Metrics (MM) approach assesses the security, privacy and dependability (SPD) triplet of a component, a sub-system and the total system.

The specific use case analysed is the privacy analysis of a medical system for diabetes measurements. The system consists of a glucometer interacting with a mobile device over Bluetooth, and a host application for mointoring of the application. The example is based on two metrics, authentication and encryption, being applied both for the application and the mobile operating system. The result of applying the MM method leads to privacy levels of the system, providing a privacy level between 40 and 70 for the application.

The proposed extension considers interconnections between components. In the envisaged use case, the interconnection is explicitly dominant for the mobile operating system (OS). The OS is surrounded by 4 components, and has 6 reachable components. The analysis using the interconnection extension of the MM approach leads to a reduced privacy level being between 28 and 50 for the Mobile OS.

VI. ACKNOWLEDGMENT

The authors would like to thanks the JU Artemis nSHIELD project for the development of the methodology, and the use cases showing the applicability of the approach. We would also like to thank Movation and the Research Council of Norway for enabling and supporting the PhD work.

REFERENCES

[1] S. Adibi, Ed., *Mobile health*. Springer, 2015, ISBN:9783319128160.
[2] P. Cipresso, S. Serino, D. Villani, C. Repetto, L. Sellitti, G. Albani, A. Mauro, A. Gaggioli, and G. Riva, "Is your phone so smart to affect your state? An exploratory study based on psychophysiological measures," *Neurocomputing*, vol. 84, pp. 23–30, 2012.
[3] J. Eloff, M. Eloff, M. Dlamini, and M. Zielinski, "Internet of People , Things and Services - The Convergence of Security, Trust and Privacy," *Proceedings of the third International CompanionAble Workshop IoPTS, Brussels*, 2009.
[4] I. Garitano, S. Fayyad, and J. Noll, "Multi-metrics approach for security, privacy and dependability in embedded systems," *Wireless Personal Communications*, vol. 81, no. 4, pp. 1359–1376, 2015.
[5] J. Noll, I. Garitano, S. Fayyad, E. Åsberg, and H. Abie, "Measurable Security, Privacy and Dependability in Smart Grids," *Journal of Cyber Security and Mobility*, vol. 3, no. 4, pp. 371–398, 2015.
[6] P. K. Manadhata and J. M. Wing, "An attack surface metric," *IEEE Transactions on Software Engineering*, vol. 37, no. 3, pp. 371–386, 2011.
[7] "National Vulnerability Data Base," 2015, URL: https://nvd.nist.gov/ [accessed: 2015-09-21].
[8] P. Ammann, D. Wijesekera, and S. Kaushik, "Scalable, graph-based network vulnerability analysis," *Proceedings of the 9th ACM Conference on Computer and Communications Security*, pp. 217–224, 2002.
[9] J. Dawkins, C. Campbell, and J. Hale, "Modeling network attacks: Extending the attack tree paradigm," *Workshop on Statistical and Machine Learning Techniques in Computer Intrusion Detection*, pp. 75–86, 2002.
[10] P. Xie, J. H. Li, X. Ou, P. Liu, and R. Levy, "Using bayesian networks for cyber security analysis," *Dependable Systems and Networks (DSN), 2010 IEEE/IFIP International Conference on*, pp. 211–220, 2010.
[11] L. Wang, T. Islam, T. Long, A. Singhal, and S. Jajodia, "An attack graph-based probabilistic security metric," *Data and applications security XXII*, pp. 283–296, 2008.
[12] L. Wang, A. Singhal, and S. Jajodia, "Measuring the overall security of network configurations using attack graphs," *Proceedings of the 21st annual IFIP WG 11.3 working conference on Data and applications security*, pp. 98–112, 2007.
[13] L. Wang, A. Liu, and S. Jajodia, "Using attack graphs for correlating, hypothesizing, and predicting intrusion alerts," *Computer communications*, vol. 29, no. 15, pp. 2917–2933, 2006.
[14] B. Schneier, "Attack trees," *Dr. Dobb's journal*, vol. 24, no. 12, pp. 21–29, 1999.
[15] R. Dantu, K. Loper, and P. Kolan, "Risk management using behavior based attack graphs," *Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on*, vol. 1, pp. 445–449, 2004.
[16] M. Howard, J. Pincus, and J. M. Wing, Eds., *Measuring relative attack surfaces*. Springer, 2005, pages:112-140.
[17] S. Fayyad and J. Noll, "Security and Safety Composition Methodology," *CENTRIC 2014*, pp. 60–65, 2014.
[18] L. Garnweidner, I. Borgen, I. n. Garitano, J. Noll, and M. Lukasse, "Designing and Developing a Mobile Smartphone Application for Women with Gestational Diabetes Mellitus Followed-Up at Diabetes Outpatient Clinics in Norway," *Healthcare*, vol. 3, no. 2, pp. 310–323, 2015.

 ISBN: 978-1-61208-440-4

Designing a Math Game for Children Using a Participatory Design Experience

I. Rodriguez, A. Puig
Applied Mathematics and Analysis
University of Barcelona
Catalonia, Spain
inmarodriguez,annapuig@ub.edu

S. Grau
UScience Tech
University of Vic
Central University of Catalonia
sergi.grau@uvic.cat

M. Escayola
Escola del Mar
Barcelona
Catalonia, Spain
montserrat.escayola@gmail.com

***Abstract*—Today teachers have in their hands learning resources, such as games, to engage primary school students in dry subjects. Nevertheless, many of these games have a mechanic very similar to in-class exercises, e.g., question -answer and match up images. Then, children end up boring. In this research, we aim to design a game for and by children. We present results of two Participatory Design (PD) sessions performed with children of 6th grade of Primary School to jointly design a game about fractions. The experience has been very fruitful both for us and for children. On one hand, they had an exciting experience as game designers, enforced their learning about fractions, and contributed with numerous and diverse ideas. On the other hand, we have co-created a game with enthusiastic stakeholders. Additionally, this research constitutes a further step in the study of methodologies and techniques of participatory game design with children. As a result of our experience we can say that participatory sessions with children should be very dynamic, with well-differentiated parts and using very different activities. We have used individual brainstorming using post-its, brainstorming and discussion in groups. We have also organised written activities, which seem to be adequate, provided that children have templates to be filled with ideas related to elements and mechanics of the game. We observed they showed really concentrated in these writing activities. Those parts of the participatory design that required the use of computers were the most appealing for them. Moreover, teachers and designers should pay attention and give children positive reinforcement regularly in order to avoid distraction during the PD sessions.**

Keywords–Serious Games;Participatory Design;child-computer Interaction.

I. INTRODUCTION

Learn by doing is already here, we can find a myriad of learning resources, which aim to provide the learner with more engaging and fruitful experiences, i.e., simulations and games [1]. Nevertheless, if we ask students to play games, they sometimes get bored and frustrated after some poor gameplay, mainly because is difficult to qualify them as ”real games” but as mere interactive quizzes.

What if we allow students to actively participate in the creation of these gaming experiences? [2][3] [4] After all, children are ideal partners for co-design mainly because they have grown with the technology and their game literacy can undoubtedly trigger the generation of creative ideas during the design process. Moreover, a psychologic study states that children today are more imaginative than those decades before [5].

Through participatory design, this research aims to bring together the worlds of designers and children in an effort to design more engaging and satisfactory games for learning. In particular, we focus on a dry math subject, fractions, which is difficult to grasp by children, as stated after some discussions with teachers of primary school of Escola del Mar in Barcelona. We thanks them for their participation.

Participatory design approaches may follow either a germinal method, where participants start the design from scratch, or a transformational one, where the design team creates an initial proposal to present to the users. In this research, we have followed the transformational method, initially presenting a basic rationale of the game concept to children and incrementally introducing them four game design tasks: characters and game settings, activities these characters may perform in the game, metres influencing both characters and environment state, and then translate learning goals into game mechanics (i.e., challenges creation).

These design tasks have been scheduled in two participatory sessions which have taken place in students' classroom. Participants of the study had already worked with fractions a few months before. Therefore, some of them could be considered ”experts” in the concept of fractions, favouring the generation of well founded and quality ideas. The results of the participatory design experience are fruitful both for us and for children. From our point of view, we have a game designed by and for children and have defined a methodology to be applied and refined in future experiences. From the perspective of children, they had a new experience as ”game designers”, and they went again over fraction concept, achieving a better grasp of their learning.

This paper is structured as follows. Section II presents the state of the art. Section III describes the steps in the participatory design. The discussion in Section IV provides a high level analysis, which results from the participatory experience. Finally, Section V concludes and gives some hints about future work.

II. RELATED WORK

Since its early introduction in the HCI field, Participatory Design (PD) has been a topic of study in the research of software and hardware design [6] [7]. Concretely, child-computer interaction has raised the attention to PD in the design of serious games for children.

In the serious game literature, several works present methods, experiences and results when involving children in all the stages of design process, with a variety of interesting results [3] [4]. Games range from those focused on emotional intelligence [8], collaborative storytelling [9], social skills [10], and social inclusion of children with disabilities [11].

A common concern raised in these research studies is the lack of children's domain content familiarity and game design literacy. Domain content familiarity depends on students back-

 ISBN: 978-1-61208-440-4

ground, and, if missing, it can be afforded either by previous basic training or by selecting user groups who adapt well to demands of PD sessions. Another key issue that is considered of great importance is to tightly couple domain content to game mechanics. In this research, we have focused on the intrinsic metaphor as introduced by Fabricatore, where player's cognitive tasks are related to the learning objectives and are relevant to achieve the goal of the game, thus being part of the activities of the game [12]. Moreover, nowadays we find a lot of children with a high degree of "game literacy", which although not being "game design literacy", can be enough valuable if participatory sessions are adequately focused and conducted.

Additionally, many of these works see the significance of the results centred in two aspects: the actual contribution of participants to the design of the game, and a very valuable domain experience gained by participants.

Related to PD methodologies, a recent research describes two case studies of a game intended to teach primary school students conflict resolution skills [2], each of them using a different PD method. In the first case study, they apply well-known germinal (i.e., generating ideas from scratch) PD methods, such as brainstorming and storyboarding, to support the generation of children' ideas in the early steps of game design. In the second one, they follow a transformational method (idea generation by modifying existing solutions) for involving children in middle steps of serious game design. Results on the former case study were not as positive as expected, mainly because of the limited domain knowledge of children. Nevertheless, children in the second case study proposed manifold and useful ideas on game mechanics and its relationships with the conflict resolution issue (learning objective). As stated before, our research follows the transformational method.

III. Stages on the design experience

We have followed the three basic stages of participatory design [13]: the initial exploration, the discovery process and the prototyping. The first stage, to know and learn from stakeholders (teachers and children), and second and third stages to elicit narratives and mechanics for the game from students, and create a lo-fi prototype of the game, respectively. Nevertheless, as will be introduced in the description of PD session 2, we did not arrive to define the prototype of the entire game, but fractions challenges which will form part of a gymkhana in the first game level.

A. Initial Exploration

We started this research work with some meetings with teachers in the school Escola del Mar in the city of Barcelona (Spain). They expressed us their concern about the difficulty on engaging students in specific math concepts. Particularly, they currently struggle on teaching the basic concept of fractions using physical objects like cords and wooden sticks, [14] and they would like to use other technical media, such as computer games. Nevertheless, available games do not fulfil teachers' expectations, mainly because they usually are really similar to those exercises performed in the blackboard, not favouring students' engagement.

We used existing fraction games to gather data (opinions, desires and feelings) from children. We met children in school labs and applied conventional contextual inquiry methods for user and task analysis, i.e., questionnaire, observation and interview.

The participants were 6 children of primary school, with ages between 10 and 12 years. Some of them had studied fractions in the previous academic year, and others had just learned the concept and basic operations with fractions.

The design team consisted of one moderator and one note taker. The study was performed individually for each student. The moderator welcomed and thanked each child for participating in the study, explained the goal of the session and what we wanted him/her to do. He also explained that the study was not an exam, but a way of getting his/her opinions and feeling about the games.

The protocol of the session consisted in three main steps: pre-gaming questionnaire, playing games, post-gaming questionnaire and short interview. In the playing game step, meanwhile the child was interacting with the game, the observer took notes. We encouraged them to think aloud. Although not all of them were able to do it, those who expressed their feelings while playing, did it pretty well.

In the pre-gaming questionnaire, we asked children general questions about their habits and likes about games: how frequently they play computer games, how many hours per week, what kind of games they like more (RPG, adventure, sport, others), how they get more fun in games (overcoming either their record or friends' scoring, living adventures, achieving levels), their preferences on play mode, either individually or multi-user. The post-gaming questionnaire included questions directly related to just played games. We asked which game they preferred and why, which one they did not like and if they thought that playing these games they would learn more about fractions. We also asked them to score the played games.

We selected four fractions games, which are representative of those we can find in the market. They differ in the metaphor, setting, story, graphics and sounds. Nevertheless, all of them share a common mechanics consisting of either questions or fractions matching.

We gave the children the following games (see Figure 1) about fractions: pipo club [15], fraction booster [16], Melvin's make a match [17] and Problem solving [18]. All these games work around fractions' basic concept and equivalent fractions.

The first game, pipo club, shows an image of the sky with three airplane parkings. The child is the pilot of the plane which must answer a question about fractions, if done correctly, he/she parks the plane in the parking, if not, the plane crashes. The second game displays a pizza and the player has to answer questions related to pizza portions. It has 5 different levels, and each level has several sublevels. The theme of the third game is witches and magic potions. There is an image of a shelf with bottles displaying fractions or figures with shaded parts. The player selects pairs of bottles/figures that represent the same fraction (equivalent fractions). In the scale displayed below, he/she can see if they match correctly. Finally, the fourth game displays a platform with a walking lizard, it stops when there is a gap. The gap is a measure (1m , 1/2 meter, etc.) and the lizard can pass through the gap only if the player correctly placed boards. These boards have different sizes and colors, depending on the fraction that they mark.

Later, we analysed the data collected during the contextual inquiry. Pre-gaming questionnaire showed that all participants play games between 2 and 5 hours per week, and they prefer adventure and role play games. When the question was related

 ISBN: 978-1-61208-440-4

to single-user versus multiuser, the majority of them prefer multi-user games, due to the competitive aspect of existing games.

After analysing the post-game questionnaire and observer notes, taken meanwhile children played the 4 fraction games, we draw on the following conclusions. Children like games with sounds or messages that encourage them to make further progress in the game, a fact corroborated by Reeves and Naas who found that computers that flatter and praise users in education software programs produce positive impact on them [19]. Children also like to have a guide or a clear indication of the target, if not they feel lost and lack interest in the game. Additionally, they do not want to answer all the time similar questions with similar format. At the time of playing they wanted to go changing between different screens, having new characters and new challenges about fractions. Some players, those who had just learned about fractions, wanted puzzles enough simple to be able to solve, but not extremely easy, some of them said "that's so easy...". When children were asked about if they got fun with the game, most of them answered "well .. not much".

The games better scored were Problem Solving (lizard) y Make-a-Match (magic portions) because students got more fun with them and seem to be less repetitive and less boring than others. Anyway, none of the games completely fulfilled children' expectative as they wanted to be immersed in a really fun experience and not in a serie of questions-answers similar to the exercices performed in the blackboard.

B. The Discovery process

In this stage, we worked with children to jointly design the fractions serious game. As game target audience is 9 to 12 years old kids, we selected a focus group of students with an intermediate level of fractions (sixth grade, with ages 10-11), rather than students of a higher grade that could add difficulty to game activities designed along PD sessions.

This stage was organised in two participatory sessions, which consisted of 25 children, the class teacher, three designers and an assistant. One of the designers actuated as moderator, introducing and explaining sessions' objectives and activities to be performed along the sessions. Session 1 was 2.5 hours long and session 2 lasted for 3.5 hours.

1) Participatory session 1: game conception: As previously mentioned, we have followed the transformational approach of participatory design. First, we introduced children the initial game concept - genre, goal, main characters, and scenario. Then, we elicited children ideas to incrementally add new elements to the game.

The group of 25 children was divided in 5 teams of 5 students, mixed in genre. The composition of each team was facilitated by the class teacher. Each team had a leader who had the role of team' spokesperson.

We presented children the initial game concept with slides (see Figure 2). 'The game scenario is a remote island in the middle of nowhere. Two shipwrecked players have to survive by eating, drinking, fighting and discovering stuffs in the island. Some of these actions have to be based on mathematical challenges related to fractions. For instance, players have to cut a part of wooden boards to build a bridge over a river that blocks the road. Game goal is to get away from the desert island.

We divided the session in three main parts. First, we asked

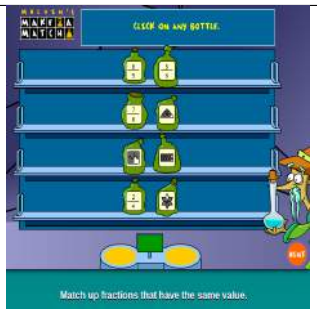
Fractions Game	Snapshot	Scoring (0-10)
PipoClub		7
Fraction Booster		7
Melvin		9
Problem Solving		9

Figure 1. Games used in contextual inquiry

them to propose objects and NPCs which may live in the island. In the second part, we asked children to think about how to escape from the island and to propose fraction-solving activities in the game. In the final part, we discussed and reviewed together emerged ideas, and we tried to obtain an initial consensus about game dynamics.

In the first part, we asked children three open questions:

1) "What kind of things do you want to find in the island?". In this case, we encouraged kids to propose stuffs that are normally in an island, but also contribute with fun and original objects of other contexts.
2) "Which characters do you want to meet in the island?". We also encourage them to be originals.

 ISBN: 978-1-61208-440-4

Figure 2. The proposed scenario: An island in the middle of nowhere and the two shipwrecked players.

Figure 3. The set of the stickers obtained from the questions about the game configuration.

3) ”If you were not a shipwrecked kid, who would you like to be?”. With this question, we wanted to know the character that children would like to be.

To gather children' answers to these questions we used the Nominal Group Technique (NGT) which is a group brainstorming technique to gather ideas but want everyone's opinions taken into account [20]. In this context, 'Nominal Group' means that people are nominally in a group but mostly work on their own. To do so, we distributed sticky notes and pens to each member of the teams [21]. Each team member wrote down the response individually on the sticky notes, promoting the participation of all team members, even shy kids. We fixed that it should be one idea per sticky note. Nevertheless, we did not fixed the ideas on a set of predefined categories, encouraging kids to be free to express their ideas. At this point, the class teacher, the three co-design experts and the design assistant acted as facilitators to help children to express their ideas.

The leader of each team collected sticky notes and filtered repeated answers. Then, he/she arranged them in a vertical cork surface (see Figure 3). A co-design expert grouped notes with similar categories and explained aloud generated ideas. We used the color-code sticky notes as team group identifier to easily visualize groups' preferences.

As a result of this first part of the session, we obtained a huge shower of ideas for the game concept. Surprisingly, some groupings of categories naturally emerged, such as vehicles, electronic devices, magical objects, sportsmen, famous people etc. Different scenarios were proposed as answers to the first question (ice-playground, soccer stadium, etc.), and a huge amount of characters were suggested in the second question (music players, sportsmen, fairies, animals, etc.). This large amount of ideas suggests us to integrate them in different levels of the game and therefore allow a variety of game scenarios. Again, in the third question about the player appeareance, kids' imagination was really far from our expectations. They proposed characters such as sailors, dancers, multimillionaires, pirates, actress, roller-bladder, fairies, survival experts, etc. As conclusion, our game design should include multiple configurations and mods to expand the game with new levels and characters.

In the second part of the session, we used brainstorming in teams of 2 students to gather design ideas about player's interactions in the game, such as how to get away from the island, how to survive, how to obtain things or how to solve a puzzle. As co-design experts, we gave to kids special scaffolds and resources to answer these questions. We gave children “templates for a design task” (i.e., a fill in the gap sentences) that allowed them to express their ideas while writting in “gaps”. This material was used in another context, to facilitate teachers the design of reading tasks [22]. For instance, we encouraged them to fill the gaps in sentences such as ”If you are a ..... and you want to get away from the island doing ....., you have to ..... in order to get .....”. Moreover, the proposed challenges had to be related to the fraction concept. The class teacher was especially important at that point to help children validate their ideas related to the fraction concept. Initially, we planned to use digital tablets to help kids to propose their ideas. However, in the live session we used paper-drawings and written activities to enhance and facilitate informal and quick interactions.

Finally, in the last part of the session the entire group shared the proposals elaborated during the brainstorming and pushed the most popular ideas under vote.

From the second and the third parts we obtained some consensus about the main game dynamics. The most preferred idea to escape the island was that both players had to construct together an engine or a ship to get away from the island. In order to obtain the pieces of the engine, players had to follow a Gymkhana through some middle challenges related to fractions (e.g., a fraction of gas is needed to get out the island, use a fraction of a rope to avoid to be killed by a carnivorous plant). Additionally, students suggested that good results could be awarded by obtaining short-cuts in the roads. Children also proposed to play some challenges competitively and some challenges in collaborative mode with the second player.

2) Participatory session 2: challenges design and lo-fi prototype: The main goal of this second session was to work with children to define the fractions' challenges that the players have to complete to achieve the goal of escaping the island.

At the beginning of the session, the moderator recapped last session. Then, she presented a video of the intro to the game. The video showed the two players, and two mysterious boxes, arriving to the island after a sinking. Each box contained some tools and appareils that could be useful for the players to survive and leave the island. These tools were some of those proposed by children in the first participatory design session, i.e., knife, iPad, pieces of wood, and rope. As seen in Figure 4 boxes' content formed part of game HUD, which is player's inventory. Note that in addition to tools, the HUD includes fractions (1/2, 3/4, 4/5) which are used to cut some proportion of rope, wood, or any other resource encountered or collected along the game.

After the video, the moderator explained the rules and dynamics of the game. The goal of the game is to escape

 ISBN: 978-1-61208-440-4

Figure 4. The proposed game HUD and the six proposed challenges.

the island. Then, when players pass each level of the game, they achieve a piece of the boat needed to go back home. During their initial stay in the island (first game level) they will participate in a gymkhana. Therefore, they have to overcome several challenges in order to arrive to a 'magic' cave. Once in the cave, they have to solve a final enigma (big boss) that conducts them to the next level of the game. To complete the challenges the player has to use resources in her inventory. Initially, each player has 5 lives which loses when she fails fraction challenges.

The participatory design session continued with a proposal of six challenges (see Figure 4) children had to select to work on:

1) The bridge challenge.
2) The oasis challenge.
3) The carnivorous plant challenge.
4) The trap challenge.
5) The treasure challenge.
6) The cave challenge.

At this point, the class was divided in 12 teams of 2 children. Each team selected one of the six challenges. Nevertheless, nobody selected the bridge challenge. Instead, they proposed a new one related to a wall.

Afterwards, and to define a challenge, they should fill an empty template with questions that should be answered by the children. In the following, we show in caps, questions children should answer. In other words, this template should capture the design of the challenge.

- SCENARIO DESCRIPTION: PART OF THE SCENARIO WHERE PLAYERS ARE SITUATED, OBJECTS IN THE SCENARIO, TOOLS IN THE INVENTORY, AND OTHERS TOOLS THEY COULD NEED TO PERFORM THE CHALLENGE.
- WHAT'S THE CHALLENGE?
- HOW CAN THE PLAYER COMPLETE THE CHALLENGE USING FRACTIONS?
  - WHICH ACTIONS SHOULD PERFORM THE PLAYERS?
  - WHICH OBJECTS DO THE PLAYER NEED? EITHER FROM THE INVENTORY OR THE SCENARIO.
  - HOW PLAYERS COULD COLLABORATE IN THIS CHALLENGE?
- WHAT HAPPENS WHETHER:
  - THE PLAYERS COMPLETE THE CHALLENGE. THAT MEANS, WHAT HAPPENS IN THE SCENARIO AND HOW IT MODIFIES PLAYERS' STATE AND INVENTORY.
  - THE PLAYERS DO NOT COMPLETE THE CHALLENGE BECAUSE:
    - THEY EITHER FAILED COMPLETELY.
    - THEY FAIL PARTIALLY AND HAS NO TOOL TO CUTS FRACTIONS OF HERBS PROPERLY.

In the following we show the example (filled) template which we gave to them, 'The roads challenge'. Again, in caps, the questions children should answer and, in italic, examples responses.

- SCENARIO DESCRIPTION: PART OF THE SCENARIO WHERE PLAYERS ARE SITUATED, OBJECTS IN THE SCENARIO, TOOLS IN THE INVENTORY, AND OTHERS TOOLS THEY COULD NEED TO PERFORM THE CHALLENGE.
  The players are in front of several roads. Some roads are closed for herbs/flowers. There is an open path that leads to the river, the player can not cross the river because there is no bridge. On the other side of closed roads, there are dangerous animals (bears and monkeys), rope and wood. The player has knives in her inventory of different sizes (fractions) that can be used for cutting vegetation.
- WHAT IS THE CHALLENGE?
  Open the roads
- HOW CAN THE PLAYER COMPLETE THE CHALLENGE USING FRACTIONS?
  - WHICH ACTIONS SHOULD PERFORM THE PLAYERS?
    Cutting the fraction of herbs that close the roads that leads to some ropes and woods.
  - WHICH OBJECTS DO THE PLAYER NEED? EITHER FROM THE INVENTORY OR THE SCENARIO.
    Knife, with the adequate size, e.g., 2/3
  - How players could collaborate in this challenge?
    The player, who has the knife with the adequate measure, cuts the herbs.
- WHAT HAPPENS WHETHER:
  - THE PLAYERS COMPLETE THE CHALLENGE. THAT MEANS, WHAT HAPPENS IN THE SCENARIO AND HOW IT MODIFIES PLAYERS' STATE AND INVENTORY.
    Both roads open and players can arrive to the roads and woods, which are added to their inventories. These materials may be useful in the next challenge, to build the bridge.
  - The players do not complete the challenge because:
    - THEY EITHER FAILED COMPLETELY.
      The bear, situated behind the herbs, eats both players.
    - THEY FAIL PARTIALLY AND HAS NO TOOL TO CUTS FRACTIONS OF HERBS PROPERLY. *Players fail consciously to die, lose one live and start again.*

 ISBN: 978-1-61208-440-4

Figure 5. The basic material delivered to kids in a power point presentation.

Figure 6. Children working in couples to create the challenges' comicboards.

Children had 30 minutes to discuss in group the definition of the challenge they had selected. Once they filled the challenge description, they had to visually describe their proposal and so produce a lo-fi prototype of the design. They used a powerpoint document to compose the scenario using different images. We provided children with a set of images that could use (See Figure 5), but they were also free to search for other images in the Internet.

This strategy is based on the Comicboarding technique, that is a variation of storyboarding intended for children that has been used to brainstorm with children aged 6 to 13. Comicboarding can be used early in the design process to depict user interactions and to capture user scenarios, cases, and tasks [23]. Figure 6 shows teams of children working in a Comicboard. The three co-design experts, the design assistant and the teacher helped the students in the process of finding the needed images that better expressed their ideas.

Figure 7 shows one of the challenges designed by a team of two children. The top-left of the figure shows the start of the challenge, the top-right the proposed fraction's challenge to be solved, the bottom-left what happens when the player hits the fence, and the bottom-right what happens when the player fails the challenge. Specifically, the description of one of the challenges designed by children was: 'Our challenge is to obtain the 8 keys to open the treasure's chest and overcome the enigma. There is a treasure's chest in front of you, but you can not open it. There are 8 keys to open the chest. The players need to find the 8 keys inside the forest to be able to open the treasure. Each key is located in a tree. When the chest is opened by the 8 keys, a fairy will appear from the chest and it will ask the next enigma to the players: Which fraction of the trees in the forest have you visited to obtain the keys? The players have to answer correctly to open the door of the cave (the next challenge).'

Figure 7. Children design.

In the template card, the children had to explain also what happens when the players perform different actions. In the same example, the children explained the following cases: 'In the case the players do not answer correctly the fairy's enigma, the fence will be opened, then the bear will attack them, and they will lose 1 life each of them', 'In order to obtain the 8 keys, the players need to collaborate using their different inventory, the players can not obtain the 8 keys alone', 'In the case they can not obtain the 8 keys, they are able to skip the current challenge to return to the island, but losing the keys they have collected'. Figure 7 shows an example result of the visual description of this challenge proposed by children.

This session was initially designed to be done in 1.5 hours, but the children were so motivated, and they wanted to work one more hour on their designs. Finally, the entire session was 3.5 hours long. The results of this session were two different descriptions for each of the six challenges. Each team explained their proposal to the rest of the class, and the best two challenges were awarded. The selection of the bests ones were made by a voting of the children. Some of the ideas obtained after this session were used to define the game design document.

Although children were rather motivated during all the session, but they were even more engaged during the creation of comicboards using the computer and searching for images in the web. Undoubtedly, computers are a tool they love to use, either for playing games or for helping in design of the game. Furthermore, they improved transversal competences such as their powerpoint skills, e.g insert images, cut & paste, etc.

IV. Results Analysis

This section provides a high level analysis of relevant issues raised during our Participatory Design (PD) experience. This analysis aims to provide foundations and useful knowledge for other researchers who will engage with children in future PD experiences.

An important factor related to PD planning is the proper scheduling of sessions. From our experience we have learned that sessions should be scheduled within close time periods. Otherwise, it is needed a considerable part of the session time

 ISBN: 978-1-61208-440-4

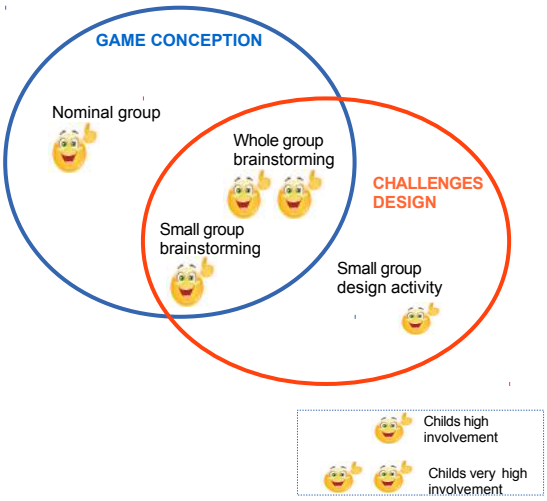


Figure 8. Childs' involvement in different phases of PD

to refresh childs' memory, and they have a harder time to put their minds in the game design endeavour.

Regarding kids' involvement, they pay more attention and are more committed and motivated in whole group brainstorming activities. However, we think that it is important to define individual activities such as nominal groups. These activities place them in context and foster an initial involvement in the design activity. Figure 8 shows kids' engagement in different activities performed during game conception and challenges design.

When it comes to facilitate children participation, both co-designers and teachers can help them to express their ideas, and material and examples are also essential to guide them in the design task. Therefore, the group of co-designers and teachers should be capable to assist all the children. We recommend at least 4 co-designers and 1 teacher for a class of 25 pupils. Additionally, when designing a participatory session with a classroom of primary school children, it should be taken into account that they need breaks and their level of attention is short. If the activity is not well defined and constrained, they may become bored, upset and a large variety of ideas may overcome the design. Designers should pay attention to give positive reinforcement at regular intervals to encourage children to continue in their work. However, once they are engaged in the problem, they may be rather creative and flexible about others' ideas.

V. Conclusions

In this paper, we have presented a participatory design experience to create a game by and for children. Particularly, this game focuses on the math concept of fractions.

We, designers, have attempted to empower children decisions about the fractions game they would like to play. Children role can be as a design-partner (early steps of the design's conception), as evaluator (with early prototypes), and as user (with the final game). We have focused in the first role, and aimed to exploit the collective creativity of students during the steps of game conception and challenges creation.

The outcomes of this research are twofold. First, the design experience constitutes a further step in the understanding and study of methodologies of participatory game design. Second, design ideas and challenges resulting from the PD sessions have been used to develop a first prototype (link to video in youtube).

As future work, we will consider a full conceptual map of fractions that children have to learn in primary school in order to consider a complete set of fractions' challenges in the game. We are also interested in providing teachers and parents with real-time monitoring of students' interactions and learning progress within the game.

Acknowledgments

Research work funded by SGR-623-2014 and TIN2012-38876-C02-02.

References

[1] C. Aldrich, The complete guide to Simulations & Serious Games, Pleiffer, Ed., 2009.

[2] R. Khaled and A. Vasalou, "Bridging serious games and participatory design," International Journal of Child-Computer Interaction, vol. 2, no. 2, 2014, pp. 93 – 100, special Issue: Learning from Failures in Game Design for Children. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2212868914000063

[3] M. Horton, J. C. Read, E. Mazzone, G. Sim, and D. Fitton, "School friendly participatory research activities with children." in CHI Extended Abstracts, J. A. Konstan, E. H. Chi, and K. Hook, Eds. ACM, 2012, pp. 2099–2104. [Online]. Available: http://dblp.uni-trier.de/db/conf/chi/chi2012a.html#HortonRMSF12

[4] A. Druin, "The role of children in the design of new technology," Behaviour and Information Technology, vol. 21, 2002, pp. 1–25.

[5] S. W. Russ and J. A. Dillon, "Changes in children's pretend play over two decades," Creativity Research Journal, vol. 23, no. 4, 2011, pp. 330 – 338.

[6] D. Schuler and A. Namioka, Eds., Participatory Design: Principles and Practices. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1993.

[7] M. J. Muller, The Human-computer Interaction Handbook. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 2003. [Online]. Available: http://dl.acm.org/citation.cfm?id=772072.772138

[8] E. Mazzone, J. C. Read, and R. Beale, "Design with and for disaffected teenagers," in Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges, ser. NordiCHI '08. New York, NY, USA: ACM, 2008, pp. 290–297. [Online]. Available: http://doi.acm.org/10.1145/1463160.1463192

[9] G. Taxen, A. Druin, C. Fast, and M. Kjellin, "Kidstory: a technology design partnership with children." Behaviour & IT, vol. 20, no. 2, 2001, pp. 119–125. [Online]. Available: http://dblp.uni-trier.de/db/journals/behaviourIT/behaviourIT20.html#TaxenDFK01

[10] J. L. Tan, D. H. Goh, R. P. Ang, and V. S. Huan, "Child-centered interaction in the design of a game for social skills intervention," Computers in Entertainment, vol. 9, 2011.

[11] B. Brederode, P. Markopoulos, M. Gielen, A. Vermeeren, and H. de Ridder, "powerball: The design of a novel mixed-reality game for children with mixed abilities," in Proceedings of the 2005 Conference on Interaction Design and Children, ser. IDC '05. New York, NY, USA: ACM, 2005, pp. 32–39. [Online]. Available: http://doi.acm.org/10.1145/1109540.1109545

[12] "Learning and videogames: an unexploited synergy," 2000, URL: http://www.learndev.org/dl/FabricatoreAECT2000.pdf [retrieved: September, 2015].

[13] E. B.-N. Sanders, E. Brandt, and T. Binder, "A framework for organizing the tools and techniques of participatory design," in Proceedings of the 11th Biennial Participatory Design Conference, ser. PDC '10. New York, NY, USA: ACM, 2010, pp. 195–198. [Online]. Available: http://doi.acm.org/10.1145/1900441.1900476

[14] M. Canals, Fracciones. Los dossiers de Maria Antonia Canals. Ed. Octaedro, ISBN: 978-8492748068, 2009.

[15] "PipoClub game," 2015, URL: http:/www.pipoclub.comjuegos-para-ninos-gratisjuego-fracciones.html [retrieved: September, 2015].

 ISBN: 978-1-61208-440-4

[16] "FractionBooster game," 2015, URL: http:/www.maths-games.orgfraction-games.html [retrieved: September, 2015].

[17] "Melvins game," 2015, URL: http:/pbskids.org/cyberchase/math-games/melvins-make-match/ [retrieved: September, 2015].

[18] "ProblemSolving," 2015, URL: http:/www.mathplayground.com [retrieved: September, 2015].

[19] B. Reeves and C. Nass, The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. New York, NY, USA: Cambridge University Press, 1996.

[20] D. Marvin, J. Campbell, and K. Jaastad, "The effect of group participation on brainstorming effectiveness for two industrial samples," Journal of Applied Psychology, vol. 47, no. 2, 2001, pp. 30–37.

[21] J. A. Fails, M. L. Guha, and A. Druin, Methods and Techniques for Involving Children in the Design of New Technology for Children. Hanover, MA, USA: Now Publishers Inc., 2013.

[22] V. Phillips and C. Wong, "Teaching to the common core by design, not accident," Kappan. Education week, vol. 93, no. 2, 2012, pp. 31–37.

[23] S. Wahid, D. S. McCrickard, J. DeGol, N. Elias, and S. Harrison, "Don't drop it!: Pick it up and storyboard," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ser. CHI '11. New York, NY, USA: ACM, 2011, pp. 1571–1580.

 ISBN: 978-1-61208-440-4

Rating Decomposition with Conjoint Analysis and Machine Learning

Florian Volk, Nadine Trüschler, Max Mühlhäuser
Telecooperation Lab
Technische Universität Darmstadt/CASED
Darmstadt, Germany
email: florian.volk@cased.de, max@informatik.tu-darmstadt.de

***Abstract*—When customers leave feedback about products, for example, a rating, they often evaluate a product as monolithic unit, neglecting that products are composed of parts with different quality, often delivered by independent suppliers. Manufacturers are more interested in individual ratings for product parts than in an overall rating. With decomposed ratings, manufacturers can improve the product quality, the selection of suppliers, and adapt pricing strategies. In this paper, we present an automated approach to decompose overall product ratings into individual ratings for product parts by the use of the results of a Conjoint analysis and supervised machine learning. Using this approach, individual ratings for product parts can be predicted with a high accuracy.**

Keywords-product ratings; decomposition; conjoint analysis; machine learning; classification; product quality; supplier selection.

I. INTRODUCTION

A recent study has shown that many customers in e-commerce scenarios are strongly influenced by feedback from other customers when making purchase decisions [1].

Customer feedback is a dually user-centric aspect of e-commerce: feedback is created by users and used by users. Product manufacturers are also interested in customer feedback as a source of quality information regarding their products. Based on customer feedback, manufacturers can identify and correct flaws in their products, as well as adapt their products to suit the expectations and requirements of customers better. Moreover, manufacturers can adapt their pricing strategy with regards to issues that are identified using feedback but cannot be corrected.

A. Composite Products

Products are usually created by a manufacturer and sold via some intermediaries to customers. However, manufacturers rarely create their products from scratch but use and assemble parts created and delivered by external third-party manufacturers, called suppliers. For example, the automobile industry heavily relies on suppliers that are independent from the car manufacturers.

The quality of a product depends on the quality of its parts, and thus, on the quality delivered by the suppliers. Manufacturers can improve the quality of their products by selecting the best known suppliers. For this, information on the individual quality of suppliers is needed.

B. Customer View on Products

Customers usually see products as monolithic units, and thus, feedback is targeting the product as a whole, but not a composite of independent parts. Manufacturers, however, are interested in obtaining individual feedback on product parts in order to learn about supplier quality.

For some types of products, customers can clearly differentiate product parts, for example, a tool and the manual found alongside the tool. In this example, low printing quality of the manual can be mentioned in the feedback. A manufacturer receiving such feedback can easily attribute the issue to the supplier of the manual.

In order to evaluate the capabilities of customers to give ratings for product parts, a study with 229 participants was conducted (29.3% female, 70.7% male, 48% university students, 44.9% employees). The participants of the study were confronted with products of increasing complexity (e.g., a nightstand, a bicycle, a hi-fi system) and were asked about their capability to identify the reasons for faults. The self-declared identification capability of the participants decreased with increasing complexity of the products. In a later task of the same study, 35.4% of the participants reported being unable to decompose their feedback into individual ratings. Another 14.8% are unsure. Additionally, the state of warranty influences customer feedback as such that faulty products are returned under warranty without the customers caring about the reason for the fault.

Generally spoken, depending on the type of product and the type of fault, it is hard to obtain decomposed ratings from customer feedback. However, manufacturers are more interested in decomposed ratings rather than in feedback on the product as a whole.

C. Contribution

In this paper, we present an automated approach to decompose overall ratings into individual ratings for suppliers. We combine results from Conjoint analyses with supervised machine learning techniques.

D. Structure

This paper is structured as follows. In Section II, we review scientific work related to our approach. Our method of combining Conjoint analyses with supervised machine learning is presented in Section III. We obtained a data set of 2,544 ratings; this data set is discussed in Section IV.

 ISBN: 978-1-61208-440-4

Section V presents and discusses the evaluation of our approach. We conclude in Section VI.

II. Related Work

In [1], Volk et al. present a decomposition-supporting review system aimed at human customers. This form-based review system is intended to assist humans in giving structured and problem-oriented feedback.

The contribution of this paper at hand does not involve user interaction during the decomposition process. The method is fully automated and operates on customer-supplied overall ratings on an ordinal scale.

A. Natural Language Processing

Most related approaches to automated feedback decomposition apply natural language processing (NLP) technologies to derive ratings or problem statements from written feedback, i.e., from reviews.

Lin and Hovy locate important sentences within paragraphs in order to create summarized reviews [2]. Instead of a summary, our proposed approach derives individual ratings.

Turney searches for keywords, as, e.g., "good" or "excellent" and then applies unsupervised learning to derive ratings on a five star scale [3] with 74% accuracy. Our approach does not need textual reviews to derive such ratings with higher accuracy.

In order to identify the reviewers' reasons for leaving feedback, Kim and Hovy assume the existence of a main statement that is either positive or negative in every review [4]. We do not share this assumption and affirm the existence of multiple main statements.

1) Decomposition by Topic

The following approaches have in common that they attempt to identify common topics within reviews. Our proposed approach is targeted at product parts, not topics, which may involve multiple product parts at once.

Dave, Lawrence, and Pennock derive (binary) sentiments towards topics [5]. Gamon et al. present a very similar approach in [6].

A summarization approach for multiple reviews is presented by Zhan, Loh, and Liu [7].

2) Decomposition by Product Features

Close to our approach is the decomposition of written reviews by product features as presented in the following publications. Product features can often be correlated with certain product parts, and, thus, associated to specific suppliers.

The feature-based summarization (FBS) system by Hu and Liu assigns binary ratings to product features derived from multiple reviews [8]. Liu, Hu, and Cheng extend the FBS system in [9]. Their *Opinion Observer* improves precision and recall over FBS.

Aciar et al. apply an ontology, which is specific to a product, to identify and rate the features of this product as described by the ontology [10]. They also calculate an overall rating. Similar, but without the need for a product-specific ontology, is the approach by Archak, Ghose, and Ipeirotis [11]. Their approach replaces the need to create ontologies first with learning product features from reviews.

By using lexicon-enhanced sentiment classification, Dang, Zhang, and Chen identify sentiments towards a product [12]. These sentiments are rated afterwards.

The closest to our approach is the *Opinion Digger* presented by Moghaddam and Ester [13]. They combine unsupervised property extraction and text mining in order to derive star ratings for product properties.

III. Method

Our approach combines the results of a traditional Conjoint analysis with supervised machine learning, more specific: with random forest classifiers [14]. For this purpose, the Conjoint analysis returns information about the importance of product aspects, or product parts. This information is fed as additional features to the machine learning process together with an overall product rating (see Figure 1).

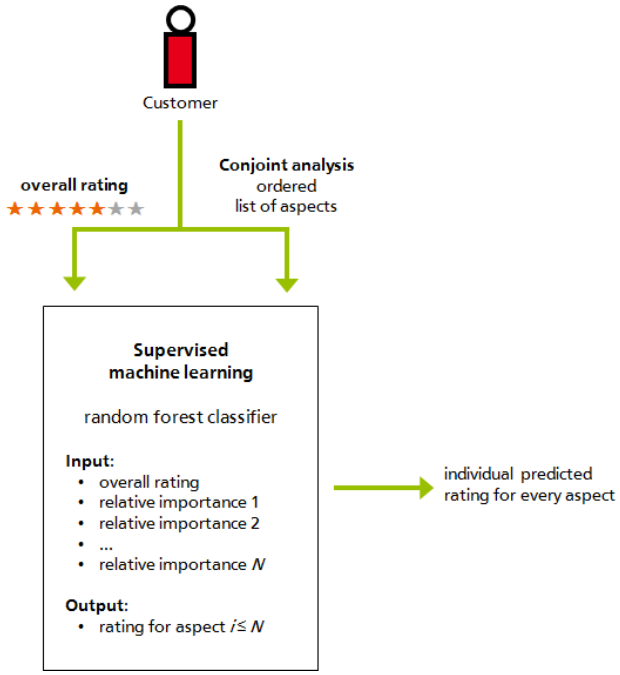


Figure 1. Combining Conjoint analyses and supervised machine learning to automatically decompose ratings.

A. Conjoint Analysis

Conjoint analyses are a set of techniques commonly used in market research. As an instance of regression analysis, Conjoint analyses identify the order of importance of product parts [15][16]. Thereby, Conjoint analyses return an importance weight for every aspect of the tested product. Conjoint analyses are often used during product design to identify the most important aspects of the product from a customer perspective. This information is used to optimally customize a product for the target market.

 ISBN: 978-1-61208-440-4

There are three main types of Conjoint analyses:

- **Traditional Conjoint analysis**. Participants of a study are asked to rate multiple instances of the same product with minimal differences for multiple aspects. Usually, only a subset of the possible combinations of stimuli (characteristics of a product aspect) is shown to the participants.
- **Adaptive Conjoint analysis**. A test system generates new combinations of stimuli based on the previous ratings by participants. This is done to focus on the product aspects that are most important for specific participants.
- **Choice-based Conjoint analysis**. The participants are presented with two products in each iteration of the analysis. Instead of rating these products, the participants only state which product is the better one according to their beliefs.

Instead of asking study participants directly for their impression of the importance for all product aspects, Conjoint analyses derive these from a multitude of products to reveal hidden relations between aspects, which the participants are not aware of.

B. Supervised Machine Learning with Random Forests

According to Witten et al. [14] and van Leeuwen [17], machine learning is the automated construction of algorithms that "learn" from data.

In supervised machine learning, a model is trained on a set of sample input data and the desired output data. The calculated model (in our case: the trained classifier) is a generalization of the input and generates its output values accordingly.

Trained models are evaluated with a second data set that only includes input data, but lacks the desired output. Comparing the generated output with the desired output (which is known to the evaluator) allows to measure the performance of the trained model.

We have compared the performance of several classifiers, namely naïve Bayes, SVM, Cart tree, and random forest classifiers. The random forest classifiers [14] (see chapter 8.3) perform best on our data set.

IV. DATA SET

Our approach was evaluated on a data set obtained from 212 human raters. Every rater participated in a traditional Conjoint analysis for two product types: digital cameras and smartphones.

The two data sets each contain 2.544 rating samples (212 raters and 12 products).

A. Participants

In a first step, the raters were confronted with different products in the form of a traditional Conjoint analysis. In this step, importance weights for the product features (assignable to product parts) were obtained. As a traditional Conjoint analysis was performed instead of a choice-based Conjoint analysis, this first step also returned overall product ratings for every tested combination of features and stimuli.

Afterwards, the raters were asked to rate all individual parts of the products. This second step generated the ratings for individual features. These ratings are used to train and evaluate the random forest classifiers.

B. Rating Scale

All products and aspects are rated on a seven-star scale ranging from one star (worst rating) to seven stars (best rating).

The seven-star scale was used instead of the five-star scale known from common internet shopping portals to enable a higher level of detail in the ratings. For the study design, we assumed that a scale of nine or more stars overstrains the differentiation capabilities of most human raters, but. We, however, wanted to achieve a higher level of granularity than a five-star scale. Furthermore, an odd number of available ratings guarantees the existence of a neutral rating.

C. Products and Stimuli

The stimuli were taken from real-world products and are given in Table I. Table I already orders the product features by their importance weights as retrieved from the Conjoint analysis.

TABLE I. PRODUCT FEATURES AND STIMULI IN THE DATA SETS

Data set	Product Feature	Stimuli
Digital cameras	Display size [inch]	2.7, 2.8, 3.0
	Zoom	3.6x, 7.1x, 40x
	Resolution [megapixels]	12, 14.2, 20
	Price [€]	310, 419, 566
Smartphones	Display size [inch]	4.0, 5.0, 5.7
	App store	Google, Apple, Microsoft
	Talking time [hours]	7, 8, 9
	Price [€]	460, 535, 777

The analysis was conducted in English and in German language. English-language participants were shown the respective real world prices in US-$ to account for influences from different tax schemes and the seller's market strategies.

V. EVALUATION

Multiple performance metrics are used to evaluate the classification performance of the trained random forest classifiers.

A. Leave-One-Out Cross Validation

A full Leave-One-Out cross validation (LpO cross validation with p=1) was conducted.

In total, 2,544 classifier models are trained on all rating samples but one. The remaining one rating sample is used to test the classifier. This way, every sample is once used to test and used 2,543 times to train the classifier. All possible combinations of samples in the data sets are used. Thus, performing an additional 10-fold cross validation was omitted.

The Leave-One-Out cross validation yields a prediction accuracy of 88%. For comparison, random guessing on a seven-star scale achieves only 14% accuracy.

 ISBN: 978-1-61208-440-4

B. *Monte-Carlo Simulation*

Full Leave-One-Out cross validation generates the most reliable performance measurements. However, training classifiers for every product feature with all available ratings (but one) is a computationally expensive task.

In order to evaluate the performance of our approach in a more realistic model, we performed a 1,000 round Monte-Carlo simulation. In every round, the classifiers are trained with just 100 randomly chosen samples from our data sets and tested on the remaining 2,444 samples. A lower number of samples reduces the computational effort for training the classifier model. The results are shown in Figure 2.

1) Compared Approaches

a) The approach proposed in this paper.
b) A baseline approach, which uses the same supervised machine learning technique, but the training set omits the importance weights derived from the Conjoint analysis.
c) Randomly guessing a rating on a seven-star scale.

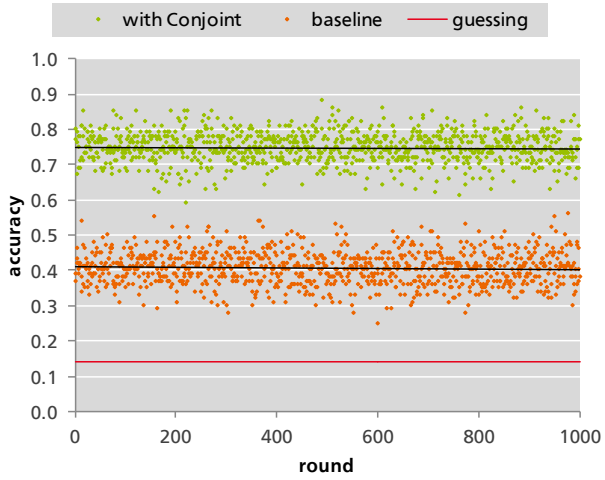


Figure 2. Comparison of the contributed approach with a baseline approach and guessing.

As can be seen, our proposed approach outperforms the other two considerably. With only 100 samples in the training set, an average accuracy of 75% is achieved. The baseline approach on average achieves slightly more than 40% accuracy. Guessing only selects the correct rating in 14% of the cases.

2) Detailed Performance Metrics

Figure 3 shows the measures performance of the trained random forest classifiers in detail. The performance is measured with standard measures for categorical data, e.g., [18].

As can be seen, the best results are achieved for the price of the smartphone and the smartphone display size. Worst performance is measured for the smartphone app store and the smartphone talk time.

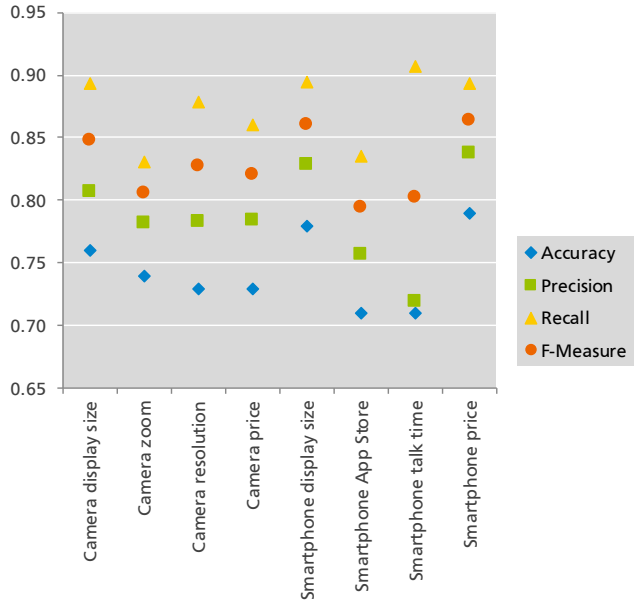


Figure 3. Detailed performance metrics of the trained random forest classifiers (average values).

C. *Accuracy in Ordinal Classification*

Even though the performance of our approach is considerably better than the compared approaches, it is even better than what the chosen performance metrics indicate.

Accuracy, precision, recall, and f-measure are calculated based on the amount of true positives, true negatives, false positives, and false negatives. In binary classification tasks, predicted ratings are classified according to the classes shown in Table II.

TABLE II. CONFUSION MATRIX FOR BINARY CLASSIFICATION TASKS

		prediction	
		true	*false*
observation	*true*	true positive	false negative
	false	false positive	true negative

However, predicting ratings is an ordinal classification task. As such, there are multiple options for a misclassification with different severity.

For example, predicting a rating of 4 stars rating instead of a (correct) 5 stars rating is less severe than predicting a 1 star rating. Nevertheless, both misclassifications affect the standard measures the same way.

While there are some metrics, to the best of our knowledge, no standard evaluation method for ordinal classification has been established, yet [19][20].

 ISBN: 978-1-61208-440-4

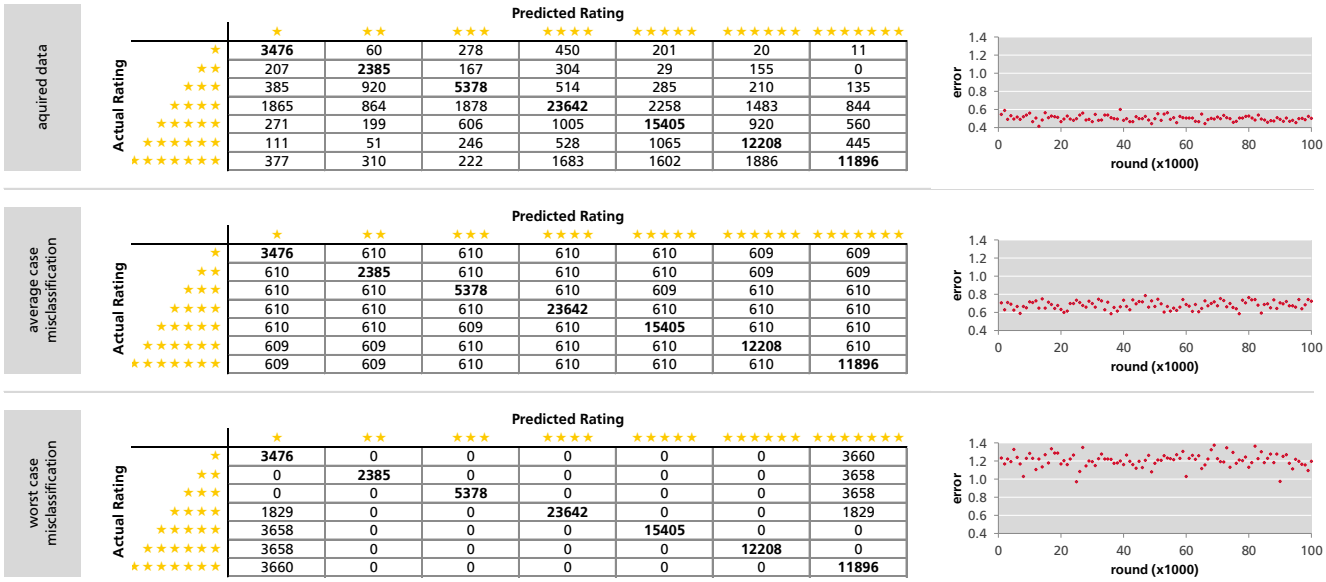


Figure 4. Confusion matrices with the same accuracy, precision, and recall, but with different prediction performance (shown as absolute prediction error).

On the example of accuracy, precision, recall, and one specific random forest classifier from the Monte-Carlo simulation, Figure 4 illustrates how our approach performs better than indicated.

- On top, the ratings predicted by the trained classifier are shown alongside with the absolute prediction error |actual rating – predicted rating|. As can be seen, the typical absolute prediction error is 0.5 stars.
- In the middle, the confusion matrix of the average misclassification case is shown, i.e. correct classifications are as returned by the trained classifier, but the misclassifications are evenly distributed over all classes. The error is about 0.7 stars. However, for this confusion matrix, the same accuracy, precision and recall are calculated as for the first case, which has predicts more correct ratings.
- The worst case misclassification case is shown in the bottom row of Figure 4. Again, all correct classification are as returned by the trained classifier. All misclassifications are maximally wrong, i.e., |actual rating – predicted rating| is maximized. Here, the error is about 1.2 stars, still yielding the same accuracy, precision and recall as all the other two cases. Additionally, the standard deviation of the prediction error (0.074 stars) is more than twice as high as for the first confusion matrix (0.032 stars).

VI. CONCLUSION

We have presented an automated approach for rating decomposition. This approach enables manufacturers to break down overall product ratings (usually given by customers) into individual ratings for product parts. Under the assumption that the product parts can be related to contributions made by external suppliers, the manufacturer is able to track the performance of its suppliers. With this information available, manufacturers can improve their choice of suppliers, and, thus, improve the quality of the products given to customers.

By supplying the results of a Conjoint analysis as additional features to a supervised machine learner, a classification model is trained, which predicts ratings for product parts based on an overall rating and importance weights for product parts.

Our evaluation shows that the proposed approach outperforms the baseline approach, which omits the results of the Conjoint analysis, as well as naïve guessing. In a leave-one-out cross validation, our approach achieves 88% prediction accuracy, i.e., the correct rating (on a seven star scale) is predicted in 88% of the cases.

A previous study has shown that many human customers are in need of assistance when giving decomposed reviews. Our approach is independent from user interaction, as only the overall rating is retrieved from the customer and automatically decomposed into individual ratings.

A. *Future work*

Our data set is specific to two product types and a seven-star rating scale. In order to evaluate the performance of the proposed approach in a general fashion, more tests are useful. This relates to both a wider set of products and to different rating scales.

When training the classifier, the leave-one-out cross validation requires the model to be trained on all available rating samples (but one), which implies extensive computational effort. The performed Monte-Carlo simulation only used 100 samples (3.9%) and achieved a lower, but still noticeably high accuracy of about 75%. It is subject to future work to find an optimal balance between training set size and prediction performance.

 ISBN: 978-1-61208-440-4

ACKNOWLEDGMENT

Some of the results discussed here will also be available in "Florian Volk, Detailing Reviews and Ratings for Trust-Enhanced Composition, PhD thesis, Technische Universität Darmstadt, 2015" *(to appear)*.

This work has been co-funded by the DFG as part of project S1 within the CRC 1119 CROSSING.

REFERENCES

[1] F. Volk, J. Pitzschel, and M. Mühlhäuser "Making the Most of Customer Product Reviews," in 7th International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services (CENTRIC), 2014, pp. 1–6.

[2] C.-Y. Lin and E. Hovy, "Identifying Topics by Position," in Proceedings of the fifth conference on Applied Natural Language Processing, 1997, pp. 283–290. Association for Computational Linguistics.

[3] P. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," in Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL), ACL '02, 2002, pp. 417–424. Association for Computational Linguistics.

[4] S.-M. Kim and E. Hovy, "Automatic Identification of Pro and Con Reasons in Online Reviews," in Proceedings of the COLING-ACL '06, 2006, pp. 483–490. Association for Computational Linguistics.

[5] K. Dave, S. Lawrence, and D. Pennock, "Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews," in Proceedings of the 12th International Conference on World Wide Web (WWW), 2003, pp. 519–528. ACM.

[6] M. Gamon, A. Aue, S. Corston-Oliver, and E. Ringger, "Pulse: Mining Customer Opinions from Free Text," in Proceedings of the 6th International Conference on Advances in Intelligent Data Analysis (IDA), IDA'05, 2005, pp. 121–132. Springer-Verlag.

[7] J. Zhan, H. T. Loh, and Y. Liu, "Gather customer concerns from online product reviews — A text summarization approach," in Expert Systems with Applications 36(2), 2009, pp. 2107– 2115.

[8] M. Hu and B. Liu, "Mining and Summarizing Customer Reviews," in Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), KDD '04, 2004, pp. 168–177. ACM.

[9] B. Liu, M. Hu, and J. Cheng, "Opinion Observer: Analyzing and Comparing Opinions on the Web," in Proceedings of the 14th International Conference on World Wide Web (WWW), WWW '05, 2055, pp. 342–351. ACM.

[10] S. Aciar, D. Zhang, S. Simoff, and J. Debenham, "Informed recommender: Basing recommendations on consumer product reviews," in IEEE Intelligent Systems 22(3), 2007, pp. 39–47.

[11] N. Archak, A. Ghose, and P. G. Ipeirotis, "Show me the Money! Deriving the Pricing Power of Product Features by Mining Consumer Reviews," in Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), KDD '07, 2007, pp. 56–65. ACM.

[12] Y. Dang, Y. Zhang, and H. Chen, "A Lexicon Enhanced Method for Sentiment Classification: An Experiment on Online Product Reviews," in IEEE Intelligent Systems, 25(4), 2010, pp. 46–53.

[13] S. Moghaddam and M. Ester, "Opinion Digger: An Unsupervised Opinion Miner from Unstructured Product Reviews, " in Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM), CIKM '10, 2010, pp. 1825–1828. ACM.

[14] I. H. Witten, E. Frank, and M. A. Hall, „Data Mining: Practical Machine Learning Tools and Techniques, Third Edition," Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2011.

[15] J. Huber, "A History of Choice-Based Conjoint," in 2001 Sawtooth Software Conference Proceedings, 2001, pp. 213–223, Sequiem, WA, USA. Sawtooth Software.

[16] J. J. Louviere and G. Woodworth, "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data," in Journal of Marketing Research 10, 1983, pp. 350–367.

[17] J. van Leeuwen, "Approaches in Machine Learning," in Philips Symposium on Intelligent Algorithms (SOIA'02), SOIA. Philips Research Laboratories, 2002.

[18] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," in Biometrics 33(1), 1977, pp. 159–174.

[19] L. Gaudette and N. Japkowicz, "Evaluation Methods for Ordinal Classification," in Advances in Artificial Intelligence, LNCS vol. 5549, 2009, pp. 207-210.

[20] E. Frank and M. Hall, "A Simple Approach to Ordinal Classification," in Machine Learning (ECML), LNCS vol. 2167, 2001, pp. 145-156.

 ISBN: 978-1-61208-440-4

Towards Participatory Methods to Take into Account Future Users and Future Usages of Hydrogen Energy: a Prospective Ergonomics Approach.

Robin Vivian, Eric Brangier, Corinne Bornet, Benoit Roussel, Javier Barcenilla, Amelie Bost

Perseus laboratory
Lorraine University
Metz, France
{firstname.name}@univ-lorraine.fr

Abstract - **The energetics transition will be a major challenge in the next years. Today, we must imagine needs and uses of alternative energy such as hydrogen. But how can we evaluate the technological acceptance of products and services where they do not yet exist? In this presentation, we introduce our theoretical approach of Prospective Cognitive Ergonomics, our methodology and the results obtained through the Hydrogen project, a research for exploring what might be the future use of this energy. For collecting data, we created three staffs of professionals (production, transport and energy distribution) and we analyzed their conversations (lexical analysis). This paper presents the entire process of this lexical analysis, which enabled us to validate four user models and to position them along two axes.**

Keywords - prospective; ergonomics; energy; user-centred concept.

I. INTRODUCTION

Trying to understand how an object (a phone, a watch, a car) or a computer application will be accepted by users is relatively simple. We can create specific groups, asking people to manipulate objects and analyse the strengths and weaknesses. What can we do when the object does not exist and when the concept is very fussy?

As Gibson stated [1], *"to see things is to see how to get about among them and what to do or not do with them"*. This seems obvious because as we move around our environment we learn and become conscious of its possibilities. The question is can we perceive things that do not exist and their possibilities.

On one hand, traditional points of view about creativity and problem solving, as gestalt theorists, assume that a new idea or creative act must be preceded by a period of incubation, before the idea appears by an unconscious mechanism called illumination. On the other hand, researches in the Ergonomics field consider that we can foster creativity and generate knowledge about the present or future use of a product by applying adequate methods, such as « scenarios based reasoning », [2], « personas creativity » [3], « analogy based approaches » [4] or « counterfactual reasoning » [5].

The objective of this work is to anticipate the hydrogen uses and identify the bottlenecks that constrain energetic transition. Since the use of hydrogen as energy is still anecdotal, it is a complex matter to require users to respond how they feel about it. This paper proposes a method called "Staffs of experts", based on the participation of professionals from three areas: production, transport and storage of hydrogen. We asked them to imagine the uses, the barriers and the behaviour of future users. After this introduction, we give the theoretical background. We describe in section 3, the method used for this study. The first results are given in Section 4. The last part is a conclusion.

II. THEORITICAL BACKGROUND: PROSPECTIVE ERGONOMICS APPROACH

Prospective ergonomics is defined as being an ergonomic intervention mode which on one hand identifies current needs, foresees changes and anticipates future users, and on the other hand, it creates systems, products or services which meet those needs according to health and safety, comfort and well-being, performance and satisfaction criteria [6]. Prospective ergonomic intervention is therefore speculative and inventive. Prospective ergonomics place the human in the central position in new practices, behaviour, and organization. Focusing on changes in human activity, the ergonomist takes on the role of a pilot who initiates, defines and manages the project from beginning to end, bringing in other actors to implement, or execute his/her plans. To achieve these aims, prospective ergonomics must fulfil three kinds of tasks:

- Identify current needs, understand their changes over the long term, anticipate future needs, define, reference, validate and asses them, put them in order of priority, and carry out simulations;
- Study technological, social, cultural, and economic factors which accelerate changes in human activity situations to which humans are required to adapt;
- Imagine, invent, create and build systems, products and services beneficial to humans, providing comfort, well-being, allowing for personal progress and a better quality of life in general.

Based on the same principles as corrective and preventive ergonomics, prospective ergonomics will complete the scope of ergonomic intervention modes by taking the future into account. This will enrich its relationship with other fields, not only psychology, biomechanics, physiology, cognitive sciences, scientific engineering, but also with sociology, anthropology, design, marketing and business management.

 ISBN: 978-1-61208-440-4

A. The part played by prospective ergonomics

Let us start from a widely shared definition (European Commission, 2002): «The prospective approach is a participatory process to design possible futures in the medium or long term, with the aim of shedding light on decisions made in the present and mobilizing the necessary means for a common commitment to action. It is above all a state of mind (anticipate and wish for) and behaviour (create and hope) in the service of present and future existence». Let us then observe all the fields involved in the prospective approach. Medicine, particularly epidemiology, has been using a prospective approach for a long time. The same goes for economics, marketing, demographics, business management, and even work organisation, etc. What motivated a prospective approach being integrated into these fields is simple and pragmatic: the prospective approach offers tools and methods which enable professionals to focus on the right questions and reduce inconsistency in reasoning for developments in the future. However, ergonomics has delayed integrating this forward looking perspective.

Our research project aims to address this delay. Applying a prospective approach in other fields will undoubtedly cause changes in our vision for the future. It broadens the perception of the problem by including demographic, economic, social, cultural, technological, historical, and legal data that highlight changes in future, trends. It keeps a watch on technologies, checks patent deposits and seeks to identify tomorrow's technologies. The prospective approach enables forecasting (often statistical but not only) related to «what will be done» from «what is being done» and «what has been done». It establishes data on changes in behaviour, consumption patterns, lifestyles, etc., and seeks to understand trends through curve models. In short, the prospective approach presents two major advantages:

- Enrichment of multi-disciplinary content: An interest in data from fields which study population, population trends and behaviour patterns (sociology, history, demographic, etc);
- Improved methods: An interest in statistics relating to large workforces over greater periods and data on the creativity of the actors involved.

B. Prospective ergonomics approach

Prospective ergonomics' aim is to develop prospective schemas by mobilizing a body of knowledge relative to the human and capable of describing needs, products, services and future systems. Through its focus and objectives, prospective ergonomics clearly differs from corrective and preventive ergonomics while being complementary. It adds little explored dimensions, including knowledge (methods, concepts and processes) that will enhance traditional approaches such as:

- Producing knowledge about future users: These are techniques that are able to understand future clients, such as the technique using personas [7][8], and scenarios. The scenario technique enables the pre-testing of new concepts and evaluation of their impact on target users.
- Stimulating creativity and organisation in innovative projects. This includes various techniques of constraint management [9] or social development of future needs [10][11] which will support user creativity.
- Previewing artefacts so as to tailor product use or future services. This involves not only making models, and prototypes, but also system simulation in immersive environments.
- Understanding forms of appropriation of artefacts, beyond the functional approach which limits usage to realizing the system's functionality. Artefacts are designed for users who in turn redesign them. Understanding this to-ing and fro-ing between the artefact and the human enables us to understand the particular sleights of hand, misuse, usage deflection, and users' creativity, which not only shows appropriation of the technology, but also sources of innovation of unprecedented usage. Exploring hitherto unseen usages is in fact an important source of innovation.
- A priori identification of errors, drop in user rates, performance reduction and other unfortunate experiences. Understanding unfortunate experiences provides suggestions for improvement while estimating performance for the future use of the artefact.
- Studying products and services successes and failures comparable to the ones we wish to create or belonging to the same technological ecosystems, in order to learn more about functionalities connectivity, mobility, user interfaces, aesthetics, etc.
- Using trends to define the direction innovation needs taking.
- Taking into account specific problems, this involves developing trend statistics, (demographics, sociology, ethnology, and economy), listing technological possibilities (patents, intellectual and industrial copyrights, development opportunities) to try and foresee market changes (sales, consumption, buying patterns). This process of putting trends, industrial and marketing data face to face serves to hone and assess prospective scenarios. In doing so, prospective ergonomics includes trends, statistical projections and market data in its reasoning, these aspects were not previously taken into consideration in the field.

III. METHOD

In the framework of the Hydrogen project, our objectives were to understand how the experts perceive the possible, the acceptable, the potential and the uses of hydrogen technology in the near or more distant future. We have used a variety of methodologies and analysis techniques.

To do so, we have adopted a step-by-step approach, where we have used and analysis techniques linked together in this order: 1-staffs of experts, 2- card sorting, 3-,thematic analysis of content 4- personas and 5- lexical automated analysis.

A community of expert staff is a group of experts [10][12] representing actors directly or indirectly involved in a community project. The experts were welcomed under the

responsibility of animators in a controlled, scripted and filmed situation.

The main goal of those staffs of experts was to collect data that could enable us to use the other methods for finally, define scenarios of future use for a technology, product or service.

The role of the animators (2 or 3 in our case) is, for example, to create a dialogue between the experts, to react on existing products and services, to express the needs, expectations and requirements of target users, and to explain useful knowledge necessary for the use.

For our study, we created three staffs of professionals on three themes: 1-Transportation (2 + 2 experts animators), 2-Producing (4 + 3 expert animators), and 3- Energy distribution (3 + 2 experts animators). Each session lasted three hours. The process of each expert staff was done in four phases:

1- Each participant's presentation,
2- Knowledge generation and discussion about needs on hydrogen and its use,
3- Presentation of a movie on the discussed theme and experts reactions/discussions on it,
4- Organization of knowledge about hydrogen and its use (card sorting method).

Discussions and interactions were filmed using the living lab of our laboratory Perseus / UL. The different groups watched videos and reacted over the content (Figure 1).

Figure 1. Powertech Hydrogen Fueling Station Overview for Fuel Cell Cars

A focus group lasted three hours. At the end of the meeting, one main question was asked:

- What is the most important ingredient for success in energy transition to hydrogen?

IV. RESULTATS AND DISCUSSION

A. *Lexical Analysis*

We have chosen to make a lexical analysis of all conversations. The objective was to find the significant words, to build thematic classes and find patterns to understand the blocking points of the acceptance of an energy transition. In addition, the analysis of the conversations was used to validate four user models and to position them along two axes.

For the lexical analysis, we started by faithfully transcribing 9 hours of recording that resulted in a text of 160 pages. After cleaning all unnecessary words, repetitions, hesitations, errors, etc., the document was still 157 pages.

To analyse the document, we used the Iramuteq software. Iramuteq stands for "Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires", in English, "interface of R for multi-dimensional text and questionnaire analysis". Iramuteq is built on top of R. This software allows analysing documents that are segmented into chunks. Input documents are plain text files that contains simple mark-up that identifies variables and topics. This allows distinguishing between:

- A text
- A text segment
- A combination of text segments

Analysing a text with Iramuteq require introducing different tags in the text. A tag is a string beginning with an * (eg *firstexpert). These tags allow cutting the corpus produced, for example, by speakers or expert.

Iramuteq analyses verbs, words, adjectives and carry out a top-down hierarchical classification. For this, it is necessary beforehand to define different parameters like number of classes or the minimal size of segment (of text). From this analysis, Iramuteq builds a dendrogram (hierarchical tree, see Figure 2).

For our study, the software generated 8 classes. The percentage for each class indicates the proportion of the class in the text. Each class contains a set of words. The first words are the most representative of the class, those that were more frequently reported.

 ISBN: 978-1-61208-440-4

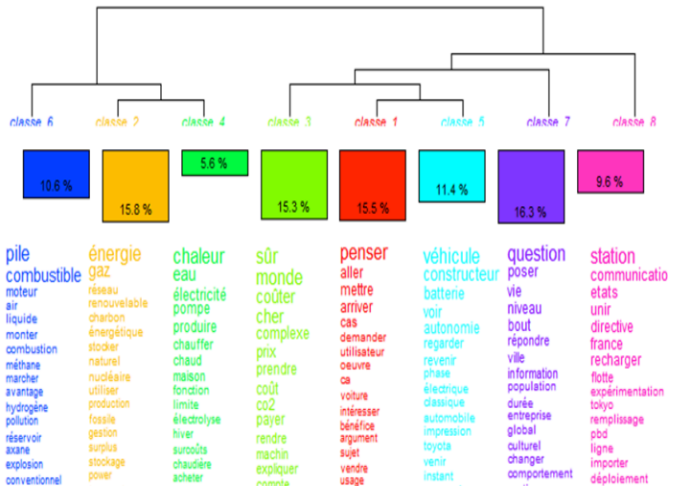


Figure 2. Our 8 classes from analysis

We can read on the first line different words like battery, energy, heat, safe, thinking, vehicle, and petrol station. The next step was to group these different classes which covered three main topics (Figure 3)

The main topics found are:

- Benefits of hydrogen and the multiplication of energies
- Choices and improvements to make in the future
- Desires of politicians and consumers

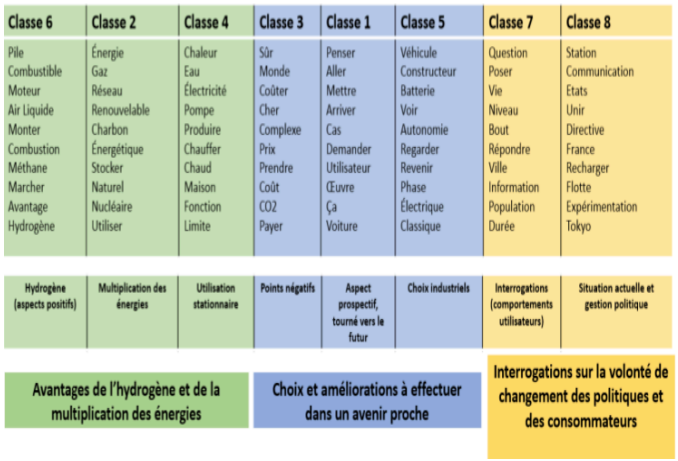


Figure 3. 8 classes grouped into three themes

In a bi-dimensional graph that resulted from a factorial analysis of morphological variables, we obtained a clear differentiation between four variables. The first two are on the X-axis (Figure 4). On the left, we can find a set of references to a social logical with words like “concern”, “charge”, “effort”, “communication”, “trust”. On the right we read more technical terms like “engine”, “fuel cell, energy, gas, heat, produce electricity”.

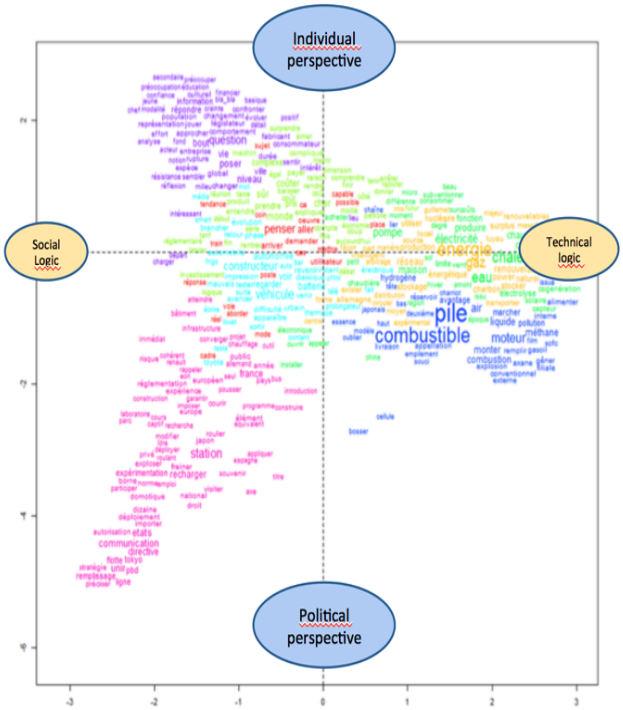


Figure 4. Factorial analysis and axes categorization of thematic content

On the Y-axis we have a comparison between what is done today in other countries and questions about what will happen in France in a more or less near future. At the bottom of the axis we find a "political concerns” with the words like "states directive, national, Tokyo, unite, etc.". At the top the preoccupation there are more individual terms like : "consumer, financial, complicated change, worry, to evolve ..."

This factorial analysis enables us to categorize future users. By placing on the axis the characteristics of a person (more individual, more social, more technical) we can find the words that characterize the potential user in terms of their expectations, needs and fears. This characterization of users can help to find for each of them the best solutions in an energy transition.

B. The elaboration of “Personas”

The concept of the personas was defined by Cooper [2] (and used by Brangier [7]) who based his facts on the notion that a user was too confusing to serve as a reference within a product team. The notion of a user was a too generic concept leading designers to develop products, which were designed for everybody but which finally, did not suit anybody. Another idea of the personas is that the necessity of adapting the product to each person will provide designers with an essential guide for a product for everybody. The persona can also serve as a communication tool, for all the stakeholders involved in the product development process.

 ISBN: 978-1-61208-440-4

The lexical analysis had two major impacts on the elaboration of user models, or personas. Firstly, the analysis enabled us to validate four profiles of users that had emerged from a thematic analysis of content

- Cécile sees hydrogen as a new constraint. For her nothing must change. She is a passive consumer
- Pierre is an active consumer. He compares, changes, manages. He likes new technologies and wants to use them to manage their energy ration.
- Laurent works in a company specialized in hydrogen. For him, the future of hydrogen depends on political decisions. For him, consumer is not important today.
- Martin's company is also specialized in energy. He thinks that hydrogen is an interesting approach but raises questions about the attitude of future consumers.

The attitudes, values and behaviors of these personas were in line with the classifications.

Secondly, personas were positioned along two axes (Figure 5): Individual perspective versus political perspective and social logic versus technical logic. Indeed, even if they share common lexical fields, each of them points out a major concern:

- Cécile : time required for energy management and safety
- Pierre : new opportunities in every day life
- Laurent : technological issues
- Martin : technological acceptance from users.

The lexical field at the lower left corner of the graph has not been purposely illustrated by the four personas. It is made up mostly of comparisons between countries, although personas, as prospective users are future orientated.

Figure 5. Positionning personas on factorial analysis graph of discourse.

Laurent is a technician. It is not concerned by the daily use difficulties. He thinks that Hydrogen is the solution to the problem of declining fossil fuels. For him, there is no alternative. He thinks the change is also a political will. Martin is a professional too. But he thinks that using hydrogen will impact the lives of people (price, constraints, etc.). For him customer perception and opinion are important. Pierre is the example of the optimist user. He enjoys new technologies. He expects to manage its energy with automation tools. This is an optimistic environmentalist. Cécile is a dubious user. She does not want changing her habits. She knows that the ecological factor is important and she wants to make an effort but she refuse to be impacted by technical constraints.

V. CONCLUSION

Our work presents a perspective design method based on lexical analysis from speech of professionals of hydrogen. We categorized into 8 classes major areas of inquiry. From classes we built a model on essential needs. We identify three major themes

- Benefits of hydrogen
- Making in the future
- Desires of politicians and consumers

The most reported words in the lists are: "battery, energy, heat, safe, thinking, vehicle, petrol station". From this, it is possible to construct a sentence that summarizes all the speech:

- "*The batteries are sources of energies of tomorrow. They can generate heat, they are safe.*"
- *"I think that vehicles in the future will be self-sufficient and hydrogen stations will be easily found".*

The "persona based" method proposes that one fictitious character can individually represent a whole category of likely future consumers (4 personas in this study). To summarise, on the one hand, the persona's features enable professionals of hydrogen to create scenarios for the use of products or services and on the other hand, they enable distributers to develop a marketing strategy for the same product or service. This work shows also that the persona based method can be enriched by a lexical analysis and it contributes to construct representations of hydrogen uses:

- In the stationary area
- In the mobility area
- In the distribution and petrol station.

This work is the beginning of a larger project. Our future works will be:

- Through a European project, we will be working on theperception of hydrogen as alternative energy. The main idea of this project will be to build information environments for specialists but also for novices. We want to develop different applications for daily users or occasional users. For example we want to build a series of online training courses for young learners (hydrogen sensitization) as well as technical training for professionals of hydrogen as the heating engineers, electricians or garage mechanics.

 ISBN: 978-1-61208-440-4

- Build Living Lab (two hydrogen pumps for cars and trucks) will study the behavior of future users.

ACKNOWLEDGMENT

Many thanks to the laboratory LEMTA (http://lemta.univ-lorraine.fr) for all the work done for this project on energy transition.

REFERENCES

[1] J. Gibson(1986). The Ecological Approach to Visual Perception, Hillsadle, Lawrence Erlbaum Associates.

[2] A. Cooper 1999. The inmates are running the asylum – Why High-Tech Products Drive Us Crazy and How to Restore the Sanity. Indianapolis: Sams Publishing.

[3] C. Bornet, E. Brangier, J. Barcenilla, PH. Deck (2013). Enrichir la créativité des ingénieurs avec l'analyse de l'activité et les personas : le cas d'un projet d'ergonomie prospective In F. Hubault (Eds). Ergonomie et Société : quelles attentes, quelles réponses ? SELF'2013, Congrès International d'Ergonomie. Paris, France.

[4] C. Tijus, & P. Brezillon, (2006). Problem solving and Creativity for Decision Support Systems. In *Creativity and Innovation in Decision Making and Decision Support*, Vol 1, F. Adam, P. Brezillon, S. Carlsson, and P. Humphreys (eds.). Ludic Publishing Ltd. 277-293.

[5] C. Tijus, J. Barcenilla, M. Rougeaux, & F. Jouen (2014). Open innovation and prospective ergonomics for smart clothes. In Marcelo Soares and Francisco Rebelo (Eds.): Advances in Ergonomics In Design, Usability & Special Populations: Part III, pp. 583-591, AHFE Conference 2014. ISBN : 978-1-4951-2108-1

[6] E. Brangier, & JM. Robert (2014). L'ergonomie prospective : fondements et enjeux. Le Travail Humain. Vol 77, n°1, 1-20.

[7] E. Brangier, & C. Bornet (2011). Persona: A method to produce representations focused on consumers' needs. In W. Karwowski, M. Soares & N. Stanton (Eds). Human Factors and ergonomics in Consumer Product Design: methods and techniques. 37-61. Taylor and Francis.

[8] E. Brangier, C. Bornet, C. Bastien, G. Michel, R. Vivian, (2012). Effets des personas et contraintes fonctionnelles sur l'idéation dans la conception d'une bibliothèque numérique. Le Travail Humain, tome 75, n°2, 121-145.

[9] N. Bonnardel (2006). Créativité et conception – Approches cognitives et ergonomiques. Marseille : Solal Editeurs.

[10] E. Brangier, J. Dinet, C. Bastien, (2009). La méthode des staffs d'experts de communautés. Orientation théorique, démarche méthodologique et application pratique. Document numérique, 12(2), 111-132.

[11] E. Brangier, J. Barcenilla, C. Bastien, J. Dinet. G. Michel. & R. Vivian (2009). The Expert Community Staff: An innovative method for capturing end-users' needs. ECCE Helsinki.

[12] E. Brangier, J. Dinet, L. Eilrich, (2009). The 7 basic functions of a digital library: Analysis of 14 focus groups about the usefulness of a digital library on the history of European construction. In M.J. Smith and G. Salvendy (Eds.): Human Interface

 ISBN: 978-1-61208-440-4

Making the System a Relational Partner: Users' Ascriptions in Individualization-focused Interactions with Companion-systems

Julia Krüger, Mathias Wahl, Jörg Frommer
University Clinic for Psychosomatic Medicine and Psychotherapy
Medical Faculty, Otto von Guericke University
Magdeburg, Germany
e-mail: {julia.krueger; mathias.wahl; joerg.frommer}@med.ovgu.de

***Abstract*—Whether or not a Companion-system is experienced as a confidant and empathic assistant depends on users' individual ascriptions to it. An ascription-based understanding of users' experiences in user-companion interaction is proposed, which focuses primarily on ascriptions of human-like characteristics, intentions, motivations or emotions to the system. It is examined with regard to the individualization-focused interactions pivotally required for the adaptation of a Companion-system to the user. The study is based on a Wizard of Oz experiment and subsequently conducted in-depth user interviews. By applying qualitative content analysis, four categories describing dimensions of users' ascriptions to the system were worked out (nature, capabilities, requirements, and relational offer) as well as two categories describing users' reactions based on these (adaptation work and self-disclosing behavior). The findings are discussed with regard to theories of psychology, philosophy and human-robot interaction. Moreover, two needs inherent in humans, which seem to be relevant for users' ascriptions in user-companion interaction, are referred to: the need for safety and the need to belong.**

Keywords-users' ascriptions; user experience; Companion-systems; individualization; human-computer interaction.

I. INTRODUCTION

Claims of individual-centered human-computer interaction (HCI) culminate in visions like that of technical systems adopting the "Companion-metaphor" [1]. Under terms like "relational agents", "sociable robots", "artificial companions" or "Companion-systems", these systems provide monitoring, personalized assistive and/or companionship services [2]. They shall i.a. be experienced as confidants [3] by their users and enable them to form emotional and long-term social attachments with them [4].

A. Companion-systems

Following the understanding of the German Transregional Collaborative Research Centre SFB/TRR 62, which aims at developing a Companion-technology for cognitive technical systems, Companion-systems are defined as follows: These are visionary cognitive technical systems which adapt their functionality to each individual user by considering his preferences, needs, abilities, requirements, current emotional state and situation. Hence, they represent available, reliable, cooperative and trustworthy empathic assistants providing also an emotional dimension to the interaction [5][6].

B. Users' Ascriptions form Users' Experiences

Besides a robust technical realization of these Companion-features, it seems crucial for the success and acceptance of Companion-systems that these features are *individually experienced* by their users.

It can be assumed that this individual experience is based on the user's interpretations of the implemented system characteristics and the system behavior. Following a *constructivist view* (e.g., [7]), these interpretations can be understood as individual *user's ascriptions* to the system [8][9]: The user himself 'constructs' his view on the system by (consciously, as well as mostly unconsciously) ascribing to it and experiences his ascription-based system view as 'objective reality'. Consequently, he chooses his behavior in reaction to his ascriptions, so that system- and self-related experiences are mutually influential. According to this understanding, users' ascriptions can be understood as 'interpretation foil' for individual *users' experiences* of Companion-systems. By following this perspective, the large body of user experience studies may be supplemented, which examine relationships between distinct psychological constructs and users' experiences as summarized overall evaluations [10][11].

It is proposed that in the case of Companion-systems, functional and structural ascriptions (like described in mental models, e.g., [12]) will supplement or even become secondary to *anthropomorphic ascriptions* of human-like mental states, e.g., motives, wishes, aims and feelings [8][9] (multiple examples can be found in literature, e.g., already [13] regarding ELIZA; [14] regarding robots like "Furbies" or "Tamagotchis"): Firstly, because Companion-systems will provide social cues to realize their Companion-functionality which may trigger the perception of interacting with a social actor (this assumption is in line with the *theory of social response* in HCI, e.g., [15]); and secondly, because the average user will be unable to explain and predict the system behavior on the basis of their complex construction and functioning and consequently will draw on ascribing mental states to it in order to interact effectively with them (*Intentional Stance* [16]).

 ISBN: 978-1-61208-440-4

C. Users' Ascriptions and Self-disclosure in Individualization-focused User-Companion Interaction

This study aims at examining users' ascriptions which arise in user-companion interaction (UCI). Therefore, researching initial interaction sequences seems promising. Therein, the user gets to know the system for the first time and questions like "How can I interact with this system?" or "What can I expect from the system?" arise, which are likely to imply lots of ascriptions to the system.

Initial situations in UCI will mostly focus on *gathering user information* to realize individual adaptation to the user and fully evolve Companion-functionality. In such sensible UCI, positive ascriptions to the system like willingness to pursue user's goals, sensibility and trustworthiness seem to be necessary to enable users to cooperatively disclose relevant information without feeling uncomfortable. Especially *intimate self-disclosure*, which may be necessary in usage contexts like, e.g., E-health or E-mental health, may entail high-risk information and induce feelings of vulnerability [17]. Ascriptions like malice or pursuit of dominance may induce such negative emotions and result in a decrease of cooperativeness, reactance and even communication break-ups. Hence, it can be assumed that users' ascriptions to the system have potential to *mediate users' self-disclosing behavior*.

For recommender systems [18] describe a mediation of self-disclosing behavior by users' individual interpretations of system characteristics and user experience: Objective system aspects are perceived by the user (e.g., recognition of differences in recommendation quality based on different algorithms) and make up the user experience which in turn constitutes the user's self-disclosing behavior. Regarding the so-called 'subjective system aspects', solely conscious user's evaluations of certain aspects, namely interaction usability, perceived quality and appeal of the system, are considered.

All in all, besides the large body of research in strategies for enhancing users' self-disclosure (*behavior-centered approach*; for an overview cf. [19]) still only little is known about users' experiences of and motives for self-disclosure to personalized systems [20] or experiential factors that mediate self-disclosure choices (*experience-centered approach*). An understanding of ascriptions in individualization phases may help to explain the often surprising findings in research on privacy decision making, namely that decisions on self-disclosure are made neglecting rational principles [21].

This study on users' ascriptions and their influences on users' behavior using the example of individualization-focused UCI is guided by the following *research questions*:

1. What do users ascribe to a Companion-system asking for personal and intimate user information (system-related experiences)?
2. How do users experience themselves in reaction to their individual ascriptions to it (self-related experiences)?

The paper is structured as follows: In Section II, the research approach will be described. Afterwards, the results will be presented in Section III and discussed according to theories of psychology, philosophy and human-robot interaction in Section IV. In Section V, future work on users' ascriptions in UCI will be outlined.

II. MATERIAL AND METHODS

Users' ascriptions are unobservable, highly individual and often unconscious. Thus, a qualitative research rationale was applied in order to examine them. It aims at describing individual meaning making processes and the resulting variety and contrasts of phenomena in the material rather than at quantifying and statistically generalizing their allocation in a population (e.g., [22]). The study bases on in-depth user interviews conducted subsequently to Wizard of Oz (WOz) experiments, which were analyzed interpretatively.

A. Experimental Basis

In order to do basic research on UCI, a widely standardized WOz experiment was conducted (cf. [23][24]). The simulated speech-based interactive dialog system should represent a kind of preliminary step towards visionary Companion-systems. It was represented by a male machine-like computer voice and a graphical interface on a computer screen (no agent). Some of the leading dialog design principles were continuous system-initiative and the avoidance of self-references (personal pronouns, active forms).

In the study presented here, the focus was on users' experiences of the first experimental module of the WOz experiment, called 'Initial Dialogue' (for details cf. [23][24]). Therein, personal, and even intimate user information were gathered for the purpose of simulating the individualization process of a future Companion-system: The text "Individualization and personalization for xxx" appears on the screen. The user is asked to give and spell his name whereupon it is complemented in the text (the display remains constant during the 'Initial Dialogue'). Hereafter, he is asked openly for self-introduction. The system summarizes all the information relevant for individualization and asks for revisal. The aim is to gather information about age, place of residence, profession, place of work, family, body height, clothing size and shoe size. Still missing information is requested on inquiry. Furthermore, users are asked about recent events in which they were emotional (happy and angry) as well as about hobbies. Finally, some questions concerning the use of and former experiences with technical devices (which devices are used for what purpose in everyday life; exemplification on positive and negative experiences) are asked. In case of very short answers, the system requested further elaboration.

All in all, the system design (no self-reference, speech-based interaction, introduction as personal assistant) and the design of the 'Initial Dialogue' (extensive standardization, minimal visualization, pretension of individualization purpose) seem appropriate for this study, because they provide openness for any kind of users' ascriptions.

 ISBN: 978-1-61208-440-4

B. Interview Design

Semi-structured interviews were conducted subsequent to the WOz experiments. They focused on users' subjective experiences of the simulated system and the interaction with it. The interview guide included an initial open narration stimulus: Users were asked to report freely on the experiment and tell about their thoughts and feelings concerning the interaction [8]. For this study, the so-called initial narrations evoked by this stimulus were analyzed. These were proven to be suitable to elicit spontaneous experiential expressions.

C. Sample

Following the criterion of theoretical saturation (cf. Section III-D; [25]) 31 interviews were analyzed. In order to maximize variance, interviews were chosen on the basis of a qualitative sample plan [26], which allowed to consider sample heterogeneity regarding age (18-28 and >60 years), gender and educational level [8].

D. Interview Analysis

The interviews were transcribed according to the guidelines of GAT-2 'Minimaltranskript' [27]. The initial narratives of the 31 interviews made up 410 transcribed pages. These were analyzed applying methods of summarizing qualitative content analysis [28]. The analysis process was accompanied by regular discussions in a group of qualitative researchers. It consisted of three main steps:

1) Breaking down the Text into Meaning Units

The initial narratives of at first 16 interviews (heterogeneous regarding age, gender and educational level) were divided into so-called meaning units (MUs) varying from word groups to paragraphs in length. MUs are text segments, which are understandable by themselves and contain one episode, idea or piece of information [29].

2) Qualitative Content Analysis, Initial Category System

MUs representing reflections upon the 'Initial Dialogue' were further processed. They were paraphrased, generalized and reduced using the methods of summarizing qualitative content analysis [28]. The resulting condensed MUs were grouped according to similarities and differences across all 16 cases. These groups constituted a first set of subcategories (third level categories), which in turn could be arranged into main categories representing a higher abstraction level (second level categories). According to the research questions, these second level categories were assigned to system- and self-related experiences (first level categories) (cf. Table 1). All first, second and third level categories made up an initial category system.

3) Maximization of Variance, Final Category System

In order to maximize variance of system- and self-related experiences, additional initial narratives were added. One by one these were broken down into MUs, which were summarized as described above. The reduced MUs of each initial narrative were arranged according to the initial category system. Categories were reformulated or new ones (second and third level) were created during this process if necessary. After adding another 15 initial narratives, no further substantial increase of variance regarding the revealed categories and their contents could be detected (theoretical saturation [25]). Thus, the category system was final.

III. RESULTS

The range of users' system- and self-related experiences during the 'Initial Dialogue' represented by first and second level categories is shown in Table 1. It was desisted from listing third level subcategories in favour of describing main issues represented by them in Sections III-A and III-B.

TABLE I. OVERVIEW OF THE CATEGORY SYSTEM

First level category	Second level category (No. of third level subcategories)
system-related experiences	Nature of the system: between man and machine (7)
	Capabilities of the system: between impressing and frightening (7)
	Requirements by the system: between expectable and strange (8)
	Relational offer of the system: between insensitive and recognizing (12)
self-related experiences	Adaptation work of the user: between degrading oneself and making oneself available (6)
	Self-disclosing behavior of the user: between subjection and control (11)

A. System-related Experiences

Users differ in the wealth and depth of their reflections on both, the system and themselves. Across all of them, more system-related experiences are reported.

1) Nature of the System: Between Man and Machine

Ascriptions regarding system's nature are found in 22 initial narratives. These ascriptions oscillate between the poles 'machine' and 'human-like counterpart'. Descriptions are often made by comparing the experienced interaction to human-human interaction (HHI). Thereby, system aspects, which are far away from HHI (technical sound of system's voice, lack of reciprocity and of visible emotionality) and those associated with HHI (speech-based interaction style, personal communication contents) can get in conflict with each other (*"strange to talk about personal things with such a machine"*, MH [in order to ensure anonymity, initials of each user were used]). This mismatch can result in experiencing hybrid forms of the system, which combine both, human- as well as machine-like aspects (*"it isn't a real human being"*, SS). The hybrid is experienced as unfamiliar; it causes uncertainty and even uncanniness. This can result in continually searching for the real nature of the counterpart.

2) Capabilities of the System: Between Impressing and Frightening

Ascriptions regarding system's capabilities are found in 22 initial narratives. A lot of users are impressed and astonished by the capabilities and performances of the system. Comparing the performance of the system with that of other technical systems or with human capabilities leads to the ascription of communication abilities to the system (*"it looks simple but it can do more (…) not only yes or no (…) but even that it (…) is able to communicate somewhat"*,

BH). However, there are users, who do not appraise the experienced human-like characteristics as generally positive: A system which gives the impression of a machine but shows unexpected humanly performance seems scary. Feelings of discomfort, uncertainty and uneasy skepticism can appear. Furthermore, ambivalence regarding system's capabilities results, which is often related to the ascription of the ability to abuse confidence to the system.

3) Requirements by the System: Between Expectable and Strange

Ascriptions regarding the requirements by the system are found in 26 initial narratives. System's requests are mostly experienced as strange, unexpected and surprising. Emotional reactions vary from amusement and curious excitement regarding the further interaction, to shock, uncertainty, scepticism, distrust and discomfort. The latter are associated with the missing revelation of assumed aims of the system, which (possibly even on purpose) are expected to lie hidden in system's requests (*"there I'll be kept in the dark completely"*, SP). Furthermore, uncertainty emerges from doubts regarding the meaningfulness of system's demands (*"as it asked me I thought (…) that it really cobbles something individual-specific and when I see this retrospectively, I don't know why it needed this"*, CT) or ambiguity about system's expectations regarding content, demonstration and extent of user's answers. Nevertheless, system's requests are sometimes also described as ordinary or expectable (*"some kind of standard questions so I knew something like that would happen"*, TB).

4) Relational Offer of the System: Between Insensitive and Recognizing

Ascriptions regarding the relational offer of the system are found in 22 initial narratives. On the one hand, most users describe the system's information collection behavior as negative and insensitive: The system is experienced as nosy, it exceeds limits of privacy and intimacy and applies pressure and compulsion in terms of subjection to get user information (*"so if I didn't give the required information (…) it sort of refused to go on (…) there I felt very constrained (…) yes a little blackmail"*, SP). Thereby, it is not honestly interested in the user. Some users feel tested and assessed by the system, e.g., regarding intelligence. On the other hand, there are users experiencing the system as a supporting one, which tries to adapt to them. It is experienced as being really interested in personally recognizing the user and even represents a relational partner to one of the users (*"at some point you felt like someone is really interested in you (…) then little by little you built up such a bonding"*, FK).

B. Self-related Experiences

Besides reflections on self-disclosure choices, reflections on users' general contact to the system and implicit fundamental behavioral choices regarding the reaction to it appear.

1) Adaptation Work of the User: Between Degrading Oneself and Making Oneself Available

Reflections upon adaptation towards the system are found in 18 initial narratives. The users search for an adequate reaction to the nature, requirements, capabilities and relational offer ascribed to the system. Here, a global focus of adapting to the system to ensure a successful interaction can be made out. There are users who doubt or even accuse themselves of deficient cognitive abilities or failures in anticipating system's capabilities (*"you're really stupid, you just could've told everything first"*, GA) resulting in an insufficient adaptation to the system and an unsuccessful interaction in turn. Others describe to (passively) get accustomed to the specific characteristics of the system or (more actively) invest cognitive and time resources for this purpose. Finally, some go to great lengths to adapt to the system because they anticipated or recognized system deficits. They make their own capabilities available to the system and thereby support it to ensure a successful interaction, e.g., by adapting content and way of answering to the anticipated system's capabilities (*"first of all I thought what I could tell, like what understands this machine"*, UK).

2) Self-Disclosing Behavior of the User: Between Subjection and Control

Reflections upon self-disclosing behavior in the interaction are found in 26 initial narratives. Self-disclosing experiences range between conscious or even unconscious subjection and feeling independent from the system, i.e., having control over self-disclosure. There are users who do not reflect upon their disclosure (*"I've just told this (…) and I actually didn't think about it"*, AS). They work off the system's questions obediently or answer them despite doubting their meaningfulness. Others answer reluctantly, e.g., because they feel pushed, feel personal limits exceeded or experience nonspecific aims of the system (*„I found it hard to give this information because I didn't really recognize the meaning"*, SP). A few users reduce breadth and depth of their answers, e.g., because of distrust regarding possible information abuse or absent familiarity with the system. In contrast, others answer unhesitatingly, e.g., motivated by the imagination of the system as a pleasant long-term dialog partner (*"well, I could have gabbed with it"*, SB).

IV. DISCUSSION

In the following, the findings are compiled and discussed according to theories from diverse disciplines. Two human needs, which seem to fuel users' ascriptions in UCI are referred to: the need for safety and the need to belong.

A. Need for Safety as a Motivation for Users' Ascriptions

Users' emotional reactions regarding the experience of the system mainly range from uncertainty, discomfort and irritation to skepticism, fear, feelings of strangeness and weirdness. The experienced misfit and ambiguity of ascriptions regarding different system aspects seem to be the reason: The system is experienced as an unsettling counterpart, neither human nor machine. When specifying it with regard to the interplay of system aspects, this finding seems to fit in what is described as the *Uncanny Valley effect*

 ISBN: 978-1-61208-440-4

in robotics: People's acceptance of anthropomorphic robots increases up to a certain point in which the robot is highly realistic but not perfect, than decreases abruptly (the robot 'is stuck in the Uncanny Valley'), and increases again when the robot is indistinguishable from reality [30]. The inability to merge nature, capabilities and requirements of the system into a coherent picture causes lots of negative emotions, including first and foremost uncertainty. Negative ascriptions and skepticism, discomfort and fear are fueled, when difficult relational offers by the system, e.g., pressure and enforcement of information disclosure, are experienced additionally.

Because *safety needs*, including the preference of the known over the unknown, are inherent in humans (already [31]), users strive to handle these ambiguities and *regain certainty* for interacting effectively and successful with the system. By ascribing especially human-like mental states (motives, aims, emotions, etc.) known from HHI to their unsettling counterpart, they try to make sense of its nature and behavior and turn it into a certain, predictable one. This can be understood in line with the aforementioned *Intentional Stance* [16]: The user is confronted with a highly complex system and (suffering from lack of alternative useful explanations) anthropomorphizes it to make its behavior predictable and explainable. Thereby, he constructs a common interaction situation, namely one comparable to HHI, which allows choosing confidently adequate reactions to the system. This is in line with a psychological theory of anthropomorphism [32]: People tend to see nonhuman objects as human-like ones for interacting effectively in their environment. This *'effectance motivation'* triggers access and application of anthropocentric knowledge and leads to anthropomorphism.

B. *Need to Belong as a Motivation for Users' Ascriptions and Users' Self-disclosure in UCI*

In their theoretical contribution regarding the design of robot companions, [33] consider another human need when thinking about the user interacting with such robots – the *need to belong* [34]. It describes the immanent desire in humans to establish and sustain relationships. Regarding anthropomorphism, [32] describe the wish for social contact and affiliation in humans (*'sociality motivation'*), which – besides the effectance motivation (cf. Section IV-A) – facilitates the application of anthropocentric knowledge, too.

This need to form relationships provides another possible explanation for occurring anthropomorphic ascriptions in all of the four dimensions worked out in this study: The user unconsciously aims at constructing a potential relational and therefore at least social, if not human-like partner in the system. Thus, he interacts with the system by applying an *'as-if-mode'*, as if it would be a social, human-like counterpart that is able to get in contact with him.

Results regarding users' self-related experiences underline the role of the need to belong in this study. Surprisingly, a lot of users disclose requested information cooperatively although their ascriptions to the system are mostly negative ones. They do not consider terminating the communication prematurely and often do not even think about dropping cooperativeness. Instead, they reveal information despite uncertainty, doubts, skepticism and discomfort, rate themselves rather than the system as inadequate, and adapt to anticipated system's expectations and capabilities.

This may be interpreted as the users' attempt to get and stay in contact with the system. They want to understand it as well as being understood by it. Occurring ambivalence between readily providing information (although this exceeds privacy limits) on the one hand and hesitating to provide information (because of an unsettling and unpredictable counterpart) on the other hand is decided in favor of the relationship. In order to apply the calculus perspective on privacy decision making (e.g., [35]), this could be understood as the user's acceptance of the costs of intimate self-disclosure (including endurance of ambivalence and uncertainty) for the benefit of gaining a *relational partner*.

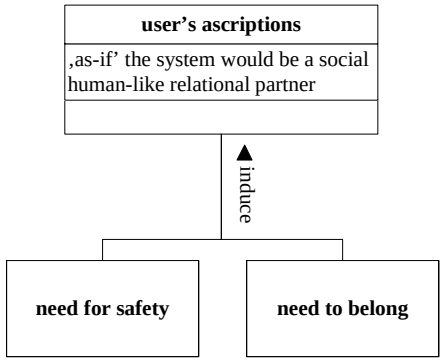


Figure 1. Grafical summary.

Figure 1 summarizes users' tendency to ascribe even human-like mental states to a Companion-system, which is based on users' inherent need for safety and need to belong.

V. CONCLUSION AND FUTURE WORK

This study examined users' individual ascriptions to a WOz-simulated speech-based interactive dialog system representing a preliminary stage of visionary Companion-systems as well as users' experiences of themselves in reaction to it. By referring to users' need for safety and need to belong, the surprisingly high user cooperativeness despite often negative ascriptions seem to be explainable. The numerous ascriptions of social skills and mental states can be interpreted as users' efforts to transform the system from an unsettling into a social, anthropomorphic counterpart representing a potential relational partner.

The presented work represents fundamental research on users' experiences in UCI. However, regarding the application-oriented perspective it can be emphasized that users tend to disclose even private and intimate information if they (independently from specific negative ascriptions to it) are enabled to see a relational partner in the Companion-system. Currently, we are conducting a supplementary study on identifying patterns in the variety of users' ascriptions and self-related experiences worked out in this study, which base

 ISBN: 978-1-61208-440-4

on the concept of 'ideal types' (cf. [36][37]). After building up this typology and assigning each individual user to exactly one of the ideal types, information about type-specific accumulations of characteristics from individual users' background profiles will be taken into account (e.g., age, sex, educational level, usage behavior and experiences regarding technical devices) to enrich the discussion of explanations for the identified patterns. The typology may be applicable to the design of Companion-systems, e.g., by deriving type-specific dialog strategies, which can foster positive ascriptions as well as reduce negative ones.

Additionally, limitations of this study indicate potentials for future research: The work was based on an experimentally simulated initial system contact for the purpose of individualization. Longitudinal studies are needed to show how users' ascriptions may change in different situations during even long term interactions. By applying quantitative and mixed methods, influences of ascriptions on observable user behavior (here with regard to self-disclosure) could be taken into account, too. Furthermore, experiments were run with volunteers. Thus, questions of ecological validity arise. In fact, motives of system use as well as downstream questions, e.g., regarding costs and benefits of self-disclosure, will become more important when considering real life UCI scenarios like, e.g., in E-Health. It could be hypothesized that being dependent on a Companion-system may increase users' tendency to ascribe in order to form a predictable, relational counterpart, as well as enhance self-disclosure in the interaction.

ACKNOWLEDGMENT

The presented study is performed in the framework of the Transregional Collaborative Research Centre SFB/TRR 62 "A Companion-Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG). The responsibility for the content of this paper lies with the authors.

REFERENCES

[1] S. Turkle, "Sociable technologies: Enhancing human performance when the computer is not a tool but a companion", in Converging Technologies for Improving Human Performance, M. C. Roco and S. B. William, Eds. Arlington: NSF-DOC report, pp. 133-140, 2002.

[2] K. Böhle and K. Bopp, "What a vision: the artificial companion. A piece of vision assessment including an expert survey", STI stud., vol. 10, 2014, pp. 155-186, ISSN: 1861-3675, Available from: http://www.sti-studies.de/ojs/index.php/sti/article/view/148/132 [retrieved: 09, 2015].

[3] Y. Wilks, "Foreword", in Close engagements with artificial companions: key social, psychological, ethical and design issues, Y. Wilks, Ed., Amsterdam: John Benjamins, pp. xi-xii, 2010.

[4] M. Pfadenhauer and C. Dukat, „Künstlich begleitet. Der Roboter als neuer bester Freund des Menschen? [Accompanied Artificially. The robot as man's new best friend?]", in Unter Mediatisierungsdruck. Änderungen und Neuerungen in heterogenen Handlungsfeldern, T. Grenz and G. Möll, Eds. Wiesbaden: Springer Fachmedien, pp. 189-210, 2014.

[5] S. Biundo and A. Wendemuth, „Von kognitiven technischen Systemen zu Companion-Systemen [From cognitive technical systems to Companion-systems]", KI, vol. 24, 2010, pp. 335-339, doi: 10.1007/s13218-010-0056-9.

[6] A. Wendemuth and S. Biundo, "A Companion Technology for Cognitive Technical Systems", Proc. of the 2011 International Conference on Cognitive Behavioural Systems (COST 2011), LNCS, 2012, pp. 89-103, doi: 10.1007/978-3-642-34584-5.

[7] E. von Glasersfeld, Radikaler Konstruktivismus: Ideen, Ergebnisse, Probleme [Radical Constructivism: Ideas, Results, Problems]. Frankfurt a. M.: Suhrkamp, 1997.

[8] J. Lange and J. Frommer, „Subjektives Erleben und intentionale Einstellung in Interviews zur Nutzer-Companion-Interaktion [Subjective experience and intentional stance in interviews regarding user-companion interaction]", Proc. der 41. GI-Jahrestagung (INFORMATIK 2011), LNI, 2011, p. 240, ISSN: 1617-5468, ISBN: 978-3-88579-286-4, Available from: www.user.tu-berlin.de/komm/CD/paper/060332.pdf [retrieved: 09, 2015].

[9] J. Frommer, D. Rösner, J. Lange, and M. Haase, "Giving Computers Personality? Personality in Computers is in the Eye of the User", in Coverbal synchrony in human-machine interaction, M. Rojc, N. Campbell, Eds. Boca Raton, FL: Crc Press, pp. 41- 71, 2013.

[10] M. Hassenzahl, "User Experience (UX): Towards an experiential perspective on product quality", in Proc. of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine (IHM 2008), ACM, 2008, pp. 11-15, doi: 10.1145/1512714.1512717.

[11] E. Karapanos, Modeling Users' Experiences with Interactive Systems. Berlin: Springer, 2013.

[12] G. C. van der Veer and M. del Carmen Puerta Melguizo, "Mental Models", in The Human-Computer Interaction Handbook. Fundamentals, Evolving Technologies and Emerging Applications, J. A. Jacko and A. Sears, Eds. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 52-80, 2003.

[13] J. Weizenbaum, "ELIZA – a computer programm for the study of natural language communication between man and machine". Commun. ACM, vol. 9, 1966, pp. 36-45, doi: 10.1145/365153.365168.

[14] S. Turkle, "Authenticity in the age of digital companions". Interact. Stud., vol. 8, 2007, pp. 501-517, doi: 10.1075/is.8.3.11tur.

[15] C. Nass and Y. Moon, "Machines and mindlessness: social responses to computers". J. Soc. Issues, vol. 56, 2000, pp. 81-103, doi: 10.1111/0022-4537.00153.

[16] D. C. Dennett, The Intentional Stance. Cambridge, MA: MIT Press, 1987.

[17] V. J. Derlega, S. Metts, S. Petronio, and S. T. Margulis, Self-Disclosure. Newbury Park, CA: Sage, 1993.

[18] B. P. Knijnenburg, M. C. Willemsen, Z. Gantner, Z., H. Soncu, and C. Newell, "Explaining the user experience of recommender systems". User Model. User-Adap., vol. 22, 2012, pp. 441-504, doi: 10.1007/s11257-011-9118-4.

[19] B. P. Knijnenburg and A. Kobsa, "Helping Users with Information Disclosure Decisions: Potential for Adaptation", Proc. of the Eighteenth International Conference on Intelligent User Interfaces (IUI 2013), ACM, 2013, pp. 407-416, doi: 10.1145/2449396.2449448.

[20] E. van de Garde-Perik, P. Markopoulos, B. de Ruyter, B. Eggen, and W. Ijsselsteijn, "Investigating privacy attitudes and behavior in relation to personalization". Soc. Sci. Comput. Rev., vol. 26, 2008, pp. 20-43, doi: 10.1177/0894439307307682.

[21] A. Acquisti and J. Grossklags, "What can behavioral economics teach us about privacy?", in Digital Privacy: Theory, Technologies, and Practices, A. Acquisti, S. Gritzalis,

 ISBN: 978-1-61208-440-4

C. Lambrinoudakis, and S. De Capitani di Vimercati, Eds. Boca Raton, FL: Auerbach Publications, pp. 363-377, 2008.

[22] T. Brüsemeister, Qualitative Forschung. Ein Überblick [Qualitative Research. An Overview]. Wiesbaden: VS Verlag für Sozialwissenschaften, 2008.

[23] D. Rösner, J. Frommer, R. Friesen, M. Haase, J. Lange, and M. Otto, "LAST MINUTE: a Multimodal Corpus of Speech-based User-Companion Interactions." Proc. of the Eighth International Conference on Language Resources and Evaluation (LREC 2012), ELRA, 2012, pp. 2559-2566, ISBN: 978-2-9517408-7-7, Available from: http://lrec.elra.info/proceedings/lrec2012/pdf/550_Paper.pdf [retrieved: 09, 2015].

[24] J. Frommer, D. Rösner, M. Haase, J. Lange, R Friesen, and M. Otto, Project A3: Detection and Avoidance of Failures in Dialogues. Wizard of Oz Experiment Operator's Manual. Lengerich: Pabst Science Publishers, 2012.

[25] J. Morse, "Theoretical saturation", in Encyclopedia of social science research methods, M. Lewis-Beck, A. Bryman, and T. Liao, Eds., Thousand Oaks, CA: Sage, pp. 1123-1124, 2004.

[26] U. Kelle and S. Kluge, Vom Einzelfall zum Typus [From Individual Case to Type]. Wiesbaden: VS Verlag für Sozialwissenschaften, 2010.

[27] M. Selting et al., "Gesprächsanalytisches Transkriptionssystem 2 (GAT 2)", Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion, vol. 10, 2009, pp. 353-402, ISSN: 1617-1837, Available from: http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf [retrieved: 09, 2015].

[28] P. Mayring, Qualitative content analysis. Theoretical foundations, basic procedures and software solution. Klagenfurth, 2014, Available from: http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173 [retrieved: 09, 2015].

[29] R. Tesch, Qualitative research analysis types and software tools. New York: Palmer Press, 1990.

[30] M. Mori, K. F. MacDorman, and N. Kageki, "The uncanny valley [from the field]". IEEE Robot. Autom. Mag., vol. 19, 2012, pp. 98–100, doi: 10.1109/MRA.2012.2192811.

[31] A. Maslow, "A theory of human motivation". Psychol. Rev., vol. 50, 1943, pp. 370-396, doi: 10.1037/h0054346.

[32] N. Epley, A. Waytz, and J. T. Cacioppo, "On seeing human: a three-factor theory of anthropomorphism". Psychol. Rev., vol. 114, 2007, pp. 864-886, doi: 0.1037/0033-295X.114.4.864.

[33] N. C. Krämer, S. Eimler, A. von der Pütten, and S. Payr, "Theory of companions: what can theoretical models contribute to applications and understanding of human-robot interaction?", Appl. Artif. Intell., vol. 25, 2011, pp. 474-502, doi: 10.1080/08839514.2011.587153.

[34] R. F. Baumeister and M. R. Leary, "The need to belong: desire for interpersonal attachments as a fundamental human motivation", Psychol. Bull., vol. 117, 1995, pp. 497-529, doi: 10.1037/0033-2909.117.3.497.

[35] H. J. Smith, T. Dinev, and H. Xu, "Information privacy research: an interdisciplinary review". MIS Quart., vol. 35, 2011, pp. 989-1015, ISSN: 0276-7783.

[36] M. Weber, "Objectivity in Social Science and Social Policy", in The Methodology of the Social Sciences, E. A. Shils and H. A. Finch, Eds. Glencoe, IL: Free Press, 1949, pp. 50-112.

[37] J. Brewer, "Ideal type", in The A-Z of Social Research, R. L. Miller and J. D. Brewer, Eds., London: Sage, pp. 147–148, 2003.

 ISBN: 978-1-61208-440-4

From Anger to Relief: Five Ideal Types of Users Experiencing an Affective Intervention in HCI

Mathias Wahl, Julia Krüger, Jörg Frommer
University Clinic for Psychosomatic Medicine and Psychotherapy, Medical Faculty
Otto von Guericke University
Magdeburg, Germany
e-mail: {mathias.wahl, julia.krueger, joerg.frommer}@med.ovgu.de

***Abstract*—Negative emotions, like distress, frustration or anger have been shown to impair the human-computer interaction (HCI). Previous research indicates that computers can reduce some of these negative emotional states in users by applying affective interventions. Until now, studies mainly measured the effectiveness of such interventions, but it is still poorly understood why these are effective and how users experience them. In the present study, in reaction to a critical dialogue situation of HCI, an affective intervention was given to deliberately frustrated users. Based on user interviews and by applying methods of qualitative research, users' meaning making processes regarding the affective intervention were explored and categories of experience were worked out. On the basis of these, five ideal types relating to how users experienced and conceptualized the affective intervention were constructed. The typology clarifies that users' appreciation of the intervention varied greatly between enthusiastic approval and definite rejection. This indicates that a 'one type fits all' solution is not appropriate for affective interventions.**

Keywords—intervention; user experience; qualitative research; ideal types; affective computing.

I. INTRODUCTION

Over the past decades, the focus in designing computer systems has shifted towards a more user centered perspective. Besides the consideration of users' demands, abilities and preferences, emotional states that arise during an interaction with a computer system have gained the interest of the human-computer interaction (HCI) community as well. Thus, the domain of 'Affective Computing', which is concerned with emotional communication between users and computer systems as well as with systems' ability to recognize and handle affective information, evolved [1]. Analyzing the impact of emotions seems reasonable, since these affect the way people make decisions, process information or interact with each other. Especially negative emotions, such as distress, frustration, anger, anxiety and sadness have been found to be associated with many unpleasant effects, like decreased ability to pay attention and think creatively or increased likelihood to take risks [2].

The consideration of emotions as an integral part of the human-computer interaction has led to the development of affective interfaces that respond to users' emotions. Besides text- and speech-based graphical user interfaces, especially embodied agents with affective capabilities have been utilized to enhance interactions. In a structured overview of the research into emotional simulation in agents, Beale and Creed [3] found out that simulated emotion can both enhance and hinder interactions. However, when looking at agents' ability to help deliberately frustrated users by applying affective interventions, results seem promising. Providing emotional support toward users led to the relief of negative emotional states like stress [4] or frustration [5], enhanced the problem solving capabilities [6] and improved the willingness to further participate in a given task [7]. A positive effect of interventions was found throughout, no matter how the interventions were implemented (e.g., solely by text [5][7], speech-based [6] or as embodied agents [4][5]).

Although these studies demonstrate that synthetic emotions expressed by computer systems can help frustrated users, it is still poorly understood why this is the case [3]. Their effectiveness was measured either by obtaining user perceptions through questionnaires (ratings of frustration, valence and arousal and ratings of the interaction in general), observing user behavior or analyzing user performance data. Besides these ratings on pre-formulated questionnaire items, the question of how users experienced the interactions and especially the affective interventions remains unanswered.

The aim of the present study is to explore how users experience an affective intervention, which was given to them in reaction to a critical dialogue situation of HCI and hence to better understand what makes such an approach valuable. More specifically, this study can be best considered as a piece of basic research at the intersection between HCI, Psychology and Social Science. It is not the intention to either evaluate the applied system, or to test the effectiveness of the applied intervention. Instead, the aim is to basically understand subjective meaning making processes regarding affective, emotion-oriented interventions applied by technical systems and thus, to answer questions like: What do users generally think about interventions?, What kind of wishes and worries do they have?, Under which circumstances are interventions valuable?, In which ways are users affected by interventions?

Analyzing in depth user experience of affective interventions above mere answers to predefined categories is a research desideratum. Hence, at first a structuring of the empirical phenomenon is needed. For that reason, an open qualitative research methodology, which aims at building a user typology, is applied in the present study. This way it is meant

 ISBN: 978-1-61208-440-4

to work out underlying structures of meaning and interpretation processes.

In the next part of this paper, related work concerning user typologies in HCI will be presented. In the third section, the empirical investigations will be described and in the fourth part, the constructed typology will be detailed. Findings will be discussed and in the end of this paper conclusions for future work will be drawn.

II. USER TYPOLOGIES

In HCI, user typologies are applied in a variety of ways, mostly to enhance design and innovation processes [8][9] or to categorize users according to their media usage behavior [10]. In the design process, user representations, such as stereotypes or personas [11], are used as a means of user modeling and rely primarily on designers' conceptualizations of users [9]. In assessing how people use media technologies, typologies are almost exclusively based on questionnaire surveys and quantitative analysis procedures, like cluster, factor or frequency analyses [10]. Thus, it seems that user types in HCI are mostly grounded on user behavior and have not been associated with user experience, since "little is known about different user types from a qualitative perspective" [10, p.949]. Even in the domain of 'User Experience' (UX), where qualitative research methods are already applied in half of the studies [12], user typologies grounded in experience seem to have been overlooked so far.

In the present study, the aim is to develop a typology of how users experience an affective intervention. Typologies are generally useful for structuring empirical phenomena by dividing people into distinct groups with high intern homogeneity and high extern heterogeneity. They can help to understand contexts of meaning and are thus particularly valuable in exploring user behavior and experience. More specifically, in this study, ideal types in the tradition of sociologist Max Weber will be constructed based on interview data. These ideal types can be thought of as idea-constructs that do not refer to perfect things, moral ideals or statistical averages, but rather stress certain elements common to most cases of a given phenomena on the basis of empirical data. An ideal type is an analytical construct that is "formed by the one-sided accentuation of one or more points of view" [13, p.90]. From a methodological point of view, ideal types can be positioned between empiricism and theory [14]: on the one side, they contribute to a fundamental understanding of users' inner processes, but on the other side, the types are abstracted in such a way, that they enable the inference of generalizations, which in turn makes findings valuable for other research areas as well.

III. EMPIRICAL INVESTIGATIONS

The present study builds on a widely standardized empirical experiment in which a critical dialogue situation of HCI was established. In reaction to this critical situation, an affective intervention was given to participants. Subsequent to taking part in the experiment, participants were interviewed, i.a., with regard to their experiences of the intervention.

A. Wizard of Oz Experiment

In order to simulate a computer system capable of accurate speech recognition and individualized reactions to user behavior, the empirical experiment was designed as a Wizard of Oz study [15]. The system was represented solely by a computer screen with a graphical user interface (without any interface agent) and a male machine-like sounding computer voice. The only way for the participants to interact with the system was via speech. In cooperation with the system, participants had to pack a suitcase for a holiday trip by selecting items from a catalogue depicted on a screen in front of them. At a certain point during the packing procedure, participants were informed about the actual weather conditions at the destination of their trip ('weather barrier'), which were different from what was suggested in the beginning. Therefore, participants were required to repack their suitcase under increasing time constraints, what was meant to cause feelings of stress and frustration. In reaction to this critical situation, an affective intervention was given to the participants (for a detailed description of the whole experimental design see [16]).

B. Affective Intervention

The affective intervention was designed to help participants in reflecting on their critical situation and to offer support for recovery. It consists of three consecutive components (cf. Table 1) and was given to the participants as a speech based audio output right after the weather barrier. The three intervention components were formulated by the research team consisting of psychotherapists in training and an experienced psychoanalyst and basically refer to the common factors of psychotherapy (activation of positive resources, actualization of what is to be changed, active help for coping with the problem, motivational clarification), which were formulated by Grawe [17].

TABLE I. INTERVENTION COMPONENTS AND CORRESPONDING SPEECH OUTPUTS

Intervention component	Speech output
1. Empathic understanding	"Because of an interruption in the data line the information about your destination could not be obtained earlier. Thus, your situation surprisingly changed. The items you chose suggest you had expected different weather conditions. If you had known the actual weather conditions of your destination, would you have chosen different items? I'm interested in your opinion."
2. Clarification of affect	"Did this situation also trigger any negative feelings? If so, can you describe them?"
3. Encouragement	"I hope your motivation to further contribute to this task was not affected by this too much."

C. Interviews

The interviews aimed at investigating how participants experienced the interaction with the system and how they were affected by the intervention. For the interviews, a semi-structured interview guide was used [18]. In the interview section relating to the intervention, the Interpersonal Process Recall (IPR) [19] was utilized. This method was applied as a specialized interview situation, in which participants watched a video of the intervention segment of their experiment. They were asked to remember and describe their experiences and feelings associated to that situation. Important to note is that they were prompted to explicitly distinguish between their reflections in hindsight and their experiences *in* the concrete situation seen in the video segment. IPR was meant to help participants to reminisce about those feelings and experiences related to the intervention, that would be forgotten typically. After an initial narration generating open question ("Please tell me what was going on inside of you in that situation"), several further questions relating to various aspects of the experience of the intervention (e.g., "How was it for you that the computer system was asking for your feelings?" or "Is there something you would have wished for in that situation?") were asked.

D. Sample

In total, there were 35 participants (17 female, 18 male) who took part in the empirical experiment, received the affective intervention and were interviewed subsequently. They were between 18 and 75 years old (two age groups: 18-28 and 60+) and had different educational backgrounds. The interview sections of these participants, which were dealing with their experiences of the intervention, form the data basis of the present study. By investigating such a heterogeneous sample, it was meant to grasp a wide range of experiences. This corresponds to the rationale of qualitative research, which can be seen in the maximization of variance and in the generation of hypotheses rather than in testing those.

E. Analysis Procedure

Initially, the audio recorded interview protocols were transcribed (which resulted in 232 transcript pages) and afterwards the transcripts were imported into the analysis software 'MaxQDA'. After this preparation of the data material, the actual data analysis began. This was conducted in two main steps: (1.) development of a category system and (2.) construction of ideal types. To ensure validity of the results, the analysis process was accompanied by regular discussions in a group of qualitative researchers (consensual validation). Moreover, to ensure reliability, the single steps of abstraction and interpretation where documented in such a way, that it remained verifiable what participants said and where the interpretation of the researchers began [20].

1) Development of a Category System

The first main step of analysis corresponds to '*summarizing qualitative content analysis*' [21][22]. At first, the text was broken down into meaning units (MUs), which are segments of text that contain one main idea and are understandable by themselves [23]. These MUs were then assigned to the one most suitable of five predefined themes: *(1.) experience of the context, (2.) experience of the system, (3.) experience of the relation to the system, (4.) self-related experience* and *(5.) experience of the intervention*. Next, the assigned MUs were paraphrased and 'streamlined' (anything that distracts from the main statement was deleted) [22], then compared to each other and grouped according to similarities, creating a set of subcategories. Finally, based on their commonalities, these subcategories were further grouped into main categories, which represent a higher abstraction level. Altogether these steps produced a hierarchy of 5 themes, 13 main categories and 58 subcategories (based on 481 MUs) – the category system.

2) Construction of Ideal Types

In the second main step of analysis (which corresponds to an approach for building typologies described by Kelle and Kluge [14]), theme 5: 'experience of intervention' was focused. In order to construct ideal types, the three main categories of this theme ('*characterization of intervention', 'subjective relevance of intervention'* and *'impact of intervention'*) were set as dimensions and the five subcategories each of these main categories was made from, were positioned on their respective dimension. Across dimensions, the subcategories were then interpretatively related to each other (cf. Figure 1). This way, five ideal types were constructed and finally each participant was assigned to exactly one type.

IV. RESULTS

Based on participants' characterization of the intervention and their appraisals of its relevance and impact for them personally, five ideal types were constructed. With the help of these ideal types, it can be demonstrated, in which diverging ways the intervention was experienced and which emotional reactions it evoked (the naming of the types refers to participants' main emotional reaction towards the intervention). In order to further elaborate the characterization of the ideal types (cf. Figure 1) and to clarify for whom of the participants the intervention was helpful, in the following each ideal type will be described in greater detail.

A. *'The Angered'*

In total, 7 participants (4 female) of all ages and levels of education were assigned to the ideal type named 'The Angered'.

At first, 'The Angered' enjoys interacting with the system, but the information about the changing weather conditions at the destination of the holiday trip disappoints him and he holds the system responsible for this ("*it knew for sure that it had disappointed me with this*", KM). He feels not to be seen and not to be taken seriously by the system, because this doesn't live up to its promises and even puts pressure on him by asking personal questions ("*why does it want to know my feelings now, of course everybody is angry when starting from false assumptions and then it laxly states neener-neener your in the wrong winter*", AM). As a result, 'The Angered' is insecure and suspicious and hence, instead of perceiving the intervention as helpful, it has a rather negative effect on him.

 ISBN: 978-1-61208-440-4

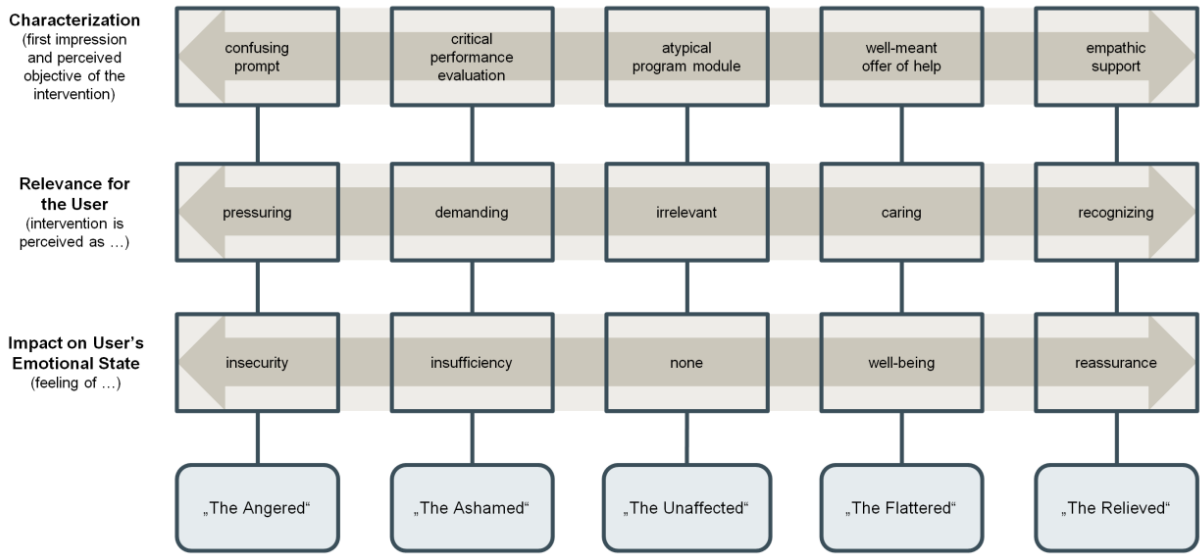


Figure 1. Dimensions of the three main categories of theme 5: 'experience of intervention'. Vertical lines indicate how the subcategories of each dimension were related to each other for the purpose of constructing the five ideal types.

B. *'The Ashamed'*

Overall, 8 participants (3 female), who all belong to the older age group and have rather higher levels of education, were assigned to the ideal type termed 'The Ashamed'.

For 'The Ashamed' it is most important to handle the experimental task in the 'right' way. He feels that the system is superior to him, that it tests him and that it critically evaluates his performance *("according to my computer at home I'm the active one and the technology is inferior, but here the technology is superior and so I have to do what it demands of me"*, WF). In his opinion, he has failed in packing the suitcase the 'right' way and therefore feelings of insufficiency arise, which are moreover strengthened by the demanding nature of the system. Still, he considers admitting his failure towards the system as his obligation, because *"when you've made a mistake, you have to take the responsibility for it"* (BP).

C. *'The Unaffected'*

Altogether, 8 predominantly young participants (2 female), with solely higher levels of education, were assigned to the ideal type named 'The Unaffected'.

'The Unaffected' considers himself as experienced with technology and therefore knows what he usually can expect from a system. To him, having an impersonal and distant working relation with the system is as important as keeping the control over the interaction (*"I've always nodded my head, because it was so obvious what it wanted at all the times, so I just nodded my approval or said no"*, FW). Generally, he doesn't want to share his feelings with a machine. Having an inappropriate suitcase for the holiday trip leaves him cold (*"I didn't have the feeling that something was going on inside of me, it was ok, bad luck"*, SK), thus he feels no need for the intervention and considers it as dispensable (*"that's not useful for me"*, HG).

D. *'The Flattered'*

Overall, 4 predominantly young participants (2 female), with mostly lower levels of education, were assigned to the ideal type termed 'The Flattered'.

'The Flattered' experiences the system as sincerely interested and for that reason he feels positive about it. At the same time he also recognizes the weaknesses of the system, but in favor of a productive collaboration he rather concentrates on its strengths (*"you can't be angry with it, it's a computer, it doesn't do anything malicious by intention, so I still felt comfortable with it"*, YD). Packing the suitcase is more of a game for 'The Flattered' and the information about the changing weather conditions rather amuses than it stresses him. He does not really need the help offered by the intervention, but he appreciates systems' empathy (*"I really liked that it was asking after me, because it was like having a cold and reserved person in front of you that suddenly offers commiserations […] that's a positive surprise"* UK).

E. *' The Relieved'*

In total, 8 participants (6 female), who are predominantly young and have mostly lower levels of education, were assigned to the ideal type named 'The Relieved'.

'The Relieved' is insecure in working on the experimental task and the weather barrier additionally stresses him *("you are thrown in at the deep end, so nobody just keeps going as if nothing has happened, of course you have lost it a little bit and you don't know what it's all about" CK)*. In his helpless state, he experiences the intervention as an empathic support that recognizes his misery (*"it asked exactly the right questions in the right moments of stress",* SS). 'The Relieved' generally feels reassured and supported by the system, which he perceives as guiding and sympathetic and as somehow human-like (*"you suspect something or someone behind it that is resonating with you"*, SD).

 ISBN: 978-1-61208-440-4

V. DISCUSSION

The aim of the present study was to understand how participants of a Wizard of Oz experiment conceptualized and experienced an affective intervention that was given to them in reaction to a critical dialogue situation. In an account of basic research, a typology was developed, which clearly demonstrates the variance of possible ways in which an intervention can be experienced. Especially for the further development of assistive technology, which aims at reacting adequately to needs and states of its users (e.g., like in the case of companion-systems [24][25]), it is important to get an idea of the far reaching impact technical support can have on users' system- as well as self-experiences.

The results presented here, indicate that users' sense making processes regarding the intervention (in terms of characterizing it and assessing its relevance and impact) are not solely positive and constructive. Besides experiencing the intervention as an offer of help ('The Relieved') or as entertaining ('The Flattered'), also characterizations as a pressuring demand ('The Angered') or a critical evaluation of performance ('The Ashamed') were found. Resulting feelings from these negative appraisals, like insecurity or insufficiency, can have a negative effect on dialogue success and the general rating of the interaction and thus present an obstacle to the actual goal of user support.

Furthermore, this study demonstrates that in case of emotional communication with a computer system, individualized support is needed. The ideal types can serve as basis for the development of type-specific interventions:

- For 'The Angered' it will be important to clearly explain how a system is functioning and which options to interact the user has.
- For 'The Ashamed' it is crucial to make clear that the system is not evaluating the user and that the user does not have to fear negative consequences.
- For "The Unaffected' it should simply be possible to turn off the intervention.
- In the cases of 'The Flattered' and 'The Relieved', the approach presented in this study seemed already pretty appropriate. In terms of a possibly more stressing event, maybe even more help (e.g., by providing concrete recommendations) could be offered.

These are only first ideas, which definitely need further elaboration and empirical validation.

The affective intervention applied in this research was intentionally kept relatively short, didn't utilize an embodied agent and was given to the participants regardless of their current emotional state. This way it was possible to gather a differentiated view on experiences, since participants were not tempted to interpret the intervention in a certain way and moreover, the meaning making processes of all kinds of participants (including those who were not in a negative emotional state and thus erroneously received the intervention, like 'The Unaffected') were included in the analysis.

In the course of building the typology, the age, sex and educational backgrounds of the participants were neglected. This way it was possible to develop the typology largely independent from pre-expectations of the researchers. However, trends in terms of these sample characteristics can be found within the ideal types, but creating user types solely on the basis of sample characteristics is not desirable. In that case, the shades between the types would get lost, the types would not be rich in substance or it would not even be possible to identify them at all. For instance, 'The Angered' is very heterogeneous in relation to the sample characteristics and would probably have been overlooked.

The sample of this study was very heterogeneous, which proved to be valuable for revealing the variance of experiences. In terms of qualitative research studies, the sample size of 35 participants is adequate [26], since it permitted a deep, case-oriented analysis that resulted in a new and richly textured understanding of experience [27]. Generally, in qualitative research it is not the intention to make predictions about the distribution of features in the population, but to reconstruct subjective meaning making processes in order to identify new aspects. Hence, not the frequency of features, but their variety is of interest [28].

The implications of the present study can primarily be seen in generating an understanding for the importance of individual meaning making processes in HCI. With the ideal types, an empirically based typology was constructed, which offers implications for argumentative generalization and theory generation. The typology reveals the 'how' and 'why' of experience as well as it illustrates the broad range of possible experiences. The types presented here, can serve as a basis for building personas, which normally rather rely on non-empiric conceptualizations of designers [11]. In working out underlying structures of meaning and exemplifying interpretation processes regarding the experience of technical support, there is also a content-related relevance for other areas of HCI, like User Experience, User Modeling or Personalization research, because these are either interested in similar content (User Experience) or can utilize the typology as a basis for adaptation (User Modeling and Personalization).

VI. CONCLUSION AND FUTURE WORK

The present study aimed at analyzing how participants experienced an affective intervention, which was given to them in reaction to a critical dialogue situation of HCI. In applying a qualitative research strategy, the experiences participants described in interviews have been structured and a user typology consisting of five ideal types has been constructed. These ideal types clarify that participants' appreciation of the intervention varied greatly between enthusiastic approval and definite rejection. This indicates that a 'one type fits all' solution is not appropriate for affective interventions. In this paper, first suggestions for the development of type specific interventions have been made, but further elaboration and empirical validation is needed.

Understanding individual experiences above relating to ratings on predefined categories or measurements of effectiveness, was the main objective of the present study. The developed ideal types clarify the importance of subjective meaning making processes: even a relatively simple intervention was capable of evoking strong feelings, like anger or shame, but also the initially intended relief. This reveals the

 ISBN: 978-1-61208-440-4

enormous potential the approach of applying affective interventions has, especially for individualized assistive technology (e.g., companion-systems). However, emotional support provided by technical systems has to be relevant for users and it should enable them to personally relate to it. Otherwise, it can have a negative effect on users' interaction experience, their overall liking of the system or in the worst case it can even lead to a decrease in cooperativeness or to communication break ups.

In the future, more empirical data is needed to further validate the described typology. It is unclear, to what extend the design of the experimental task and participants' involvement with it, as well as the design of the intervention or the demographical structure of the sample have influenced the construction of the typology. It will be interesting to explore other contexts of use or even more frustration evoking scenarios and moreover, it will be interesting to investigate another sample with a different demographical structure. However, the typology developed here can serve as a starting point for future studies with bigger sample sizes.

After further validation of the typology, it is conceivable to develop a questionnaire, with which it will be possible to quickly assign users to the appropriate user type even before an interaction begins. However, the scales such a questionnaire could be based on have to be identified first. Investigating user characteristics like attributional style, personality traits, self-efficacy or computer-experience could be a starting point for this. When it is possible to identify the user types at the beginning of an interaction, the effectiveness of type specific intervention strategies can be tested empirically.

ACKNOWLEDGMENT

The present study is performed in the framework of the Transregional Collaborative Research Centre SFB/TRR 62 "A Companion-Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG). The responsibility for the content of this paper lies with the authors.

REFERENCES

[1] R. W. Picard, Affective Computing. Cambridge, MA: MIT Press, 1997.

[2] E. Wilson, C. MacLeod, and L. Campbell, "The information-processing approach to emotion research", in Handbook of emotion elicitation and assessment, J.A. Coan and J. J. B. Allen, Eds. New York: Oxford University Press, 2007, pp. 184-202.

[3] R. Beale and C. Creed, "Affective interaction: How emotional agents affect users", Int. J. Hum-Comput. St., vol. 67, Sep. 2009, pp. 755-776, doi: 10.1016/j.ijhcs.2009.05.001.

[4] H. Prendinger, S. Mayer, J. Mori, and M. Ishizuka, "Persona effect revisited: Using bio-signals to measure and reflect the impact of character-based interfaces", The 4th International Working Conference on Intelligent Virtual Agents, 2003, pp. 283–291, doi: 10.1007/978-3-540-39396-2_48.

[5] K. Hone, "Empathic agents to reduce user frustration: The effects of varying agent characteristics", Interact. Comput. 18, vol. 18, 2006, pp. 227-245, doi: 10.1016/j.intcom.2005.05.003.

[6] T. Partala and V. Surakka, "The effects of affective interventions in human–computer interaction", Interact. Comput., vol. 16, 2004, pp. 295-309, doi: 10.1016/j.intcom.2003.12.001.

[7] J. Klein, Y. Moon, and R. W. Picard "This computer responds to user frustration: theory, design and results", Interact. Comput., vol 14, 2002, pp. 119–140, Feb. 2002, doi: 10.1016/S0953-5438(01)00053-4.

[8] D. Schuurman, K. De Moor, L. De Marez, and T. Evens, "Investigating user typologies and their relevance within a living lab-research approach for ICT-innovation", The 43rd Hawaii International Conference on System Sciences (HICSS 2010), IEEE Press, Jan. 2010, pp. 1-10, doi: 10.1109/HICSS.2010.242.

[9] T. Clemmensen, "Four approaches to user modelling—a qualitative research interview study of HCI professionals' practice", Interact. Comput., vol. 16, 2004, pp. 799-829, doi: 10.1016/j.intcom.2004.04.009.

[10] P. B. Brandtzæg, "Towards a unified Media-User Typology (MUT): A meta-analysis and review of the research literature on media-user typologies", Comput. Hum. Behav., vol. 26, Sep. 2010, pp. 940-956, doi: 10.1016/j.chb.2010.02.008.

[11] P. Turner and S. Turner, "Is stereotyping inevitable when designing with personas?", Des. Stud., vol. 32, Jan. 2011, pp. 30-44, doi: 10.1016/j.destud.2010.06.002.

[12] J. A. Bargas-Avila and K. Hornbæk, "Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience", Conference on Human Factors in Computing Systems (SIGCHI), 2011, pp. 2689-2698, doi: 10.1145/1978942.1979336

[13] M. Weber, "Objectivity in Social Science and Social Policy", in The Methodology of the Social Sciences, E. A. Shils and H. A. Finch, Eds. Glencoe, IL: Free Press, 1949, pp. 50-112.

[14] U. Kelle and S. Kluge, Vom Einzelfall zum Typus [From Individual Case to Type]. Wiesbaden: VS Verlag für Sozialwissenschaften, 2010.

[15] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, "Wizard of Oz studies—why and how", Knowl-Based. Syst., vol. 6, 1993, pp. 258-266, doi: 10.1145/169891.169968.

[16] D. Rösner, J. Frommer, R. Friesen, M. Haase, J. Lange, and M. Otto, "LAST MINUTE: a Multimodal Corpus of Speech-based User-Companion Interactions", Proc. of the Eighth International Conference on Language Resources and Evaluation (LREC 2012), ELRA, 2012, pp. 2559-2566, ISBN: 978-2-9517408-7-7, Available from: http://lrec.elra.info/proceedings/lrec2012/pdf/550_Paper.pdf [retrieved: 09, 2015].

[17] K. Grawe, "Outline of a general psychotherapy", Psychotherapeut, vol. 40, 1995, pp. 130-145, ISSN: 0935-6185.

[18] J. Lange and J. Frommer, „Subjektives Erleben und intentionale Einstellung in Interviews zur Nutzer-Companion-Interaktion [Subjective experience and intentional stance in interviews regarding user-companion interaction]", Proc. der 41. GI-Jahrestagung (INFORMATIK 2011), LNI, 2011, p. 240, ISSN: 1617-5468, ISBN: 978-3-88579-286-4, Available from: www.user.tu-berlin.de/komm/CD/paper/060332.pdf [retrieved: 09, 2015].

[19] R. Elliott, "Interpersonal Process Recall (IPR) as a psychotherapy process research method", in The psychotherapeutic process: A Research Handbook, L. Greenberg and W. Pinsof, Eds. New York, NY: Guilford, 1986, pp. 503-527.

[20] U. Flick, "Gütekriterien qualitativer Forschung [Quality criteria in qualitative research]", in Handbuch Qualitative Forschung in der Psychologie [Handbook Qualitative Research in Psychology], G. Mey and K. Mruck, Eds. VS Verlag für Sozialwissenschaften, 2010, pp. 395–407.

 ISBN: 978-1-61208-440-4

[21] P. Mayring, Qualitative content analysis. Theoretical foundations, basic procedures and software solution. Klagenfurth, 2014, Available from: http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173 [retrieved: 09, 2015].

[22] M. Schreier, Qualitative content analysis in practice. Sage Publications, 2012.

[23] R. Tesch, Qualitative research analysis types and software tools. New York: Palmer Press, 1990.

[24] Y. Wilks, Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues. Amsterdam: John Benjamins, 2010.

[25] A. Wendemuth and S. Biundo, "A Companion Technology for Cognitive Technical Systems", Proc. of the 2011 International Conference on Cognitive Behavioural Systems (COST 2011), LNCS, 2012, pp. 89-103, doi: 10.1007/978-3-642-34584-5.

[26] C. J. Yeh and A. Inman, "Qualitative data analysis and interpretation in counseling psychology: Strategies for best practices", Couns. Psychol., vol. 35, 2007, pp. 369-403, doi: 10.1177/0011000006292596.

[27] M. Sandelowski, "Sample size in qualitative research", Res. Nurs. Health, vol. 18, Apr. 1995, pp. 179-183, doi: 10.1002/nur.4770180211.

[28] K. Köhler and J. Frommer, "Criteria of sample recruiting in qualitative research", Psychother. Psychosom. Med. Psychol., vol. 61, Jun. 2011, pp. e34-5, doi: 10.1055/s-0030-1266114.

 ISBN: 978-1-61208-440-4

The Effects of Extended Estimation on Affective Attitudes in an Interactional Series of Tasks

Yoshimasa Ohmoto*, Soshi Horii* and Toyoaki Nishida*
*Department of Intelligence Science
and Technology
Graduate School of Informatics
Kyoto University
Kyoto, Japan
Email: ohmoto@i.kyoto-u.ac.jp, horii@ii.ist.i.kyoto-u.ac.jp, nishida@i.kyoto-u.ac.jp

***Abstract*—When engaged in long-term human-human interactions, we mutually estimate and construct a model of behavior. To achieve long-term, sustained human-agent interaction, the agent induces an active attitude within the human to demonstrate and share their common ground. This study aimed to investigate the effects of the agent's estimation of a human's preferences on the human's affective impressions to induce an active attitude within a series of interactions. We conducted an experiment to evaluate the effect of the proposed method using two agents. The results showed that the proposed method could reduce the number of interactions in the decision-making process and improved some of the affective impressions related to the agent's character. In addition, we found that the rate of participants' acceptance of the proposals by the agent, which was implemented our proposed method, was significantly low. This was possible because the agent provided a consistent estimation of emphasizing points for each participant. The participants' attitude indicates that they regarded the agent as being communicative.**

Keywords–Multi-modal interaction; human-agent interaction; affective attitudes.

I. INTRODUCTION

In recent years, the development of conversational agents, such as robots and virtual agents, has rapidly expanded. However, many of these agents are regarded as multimodal interfaces that provide useful information rather than as social partners [1]. There are many issues to contend with in the production of a social partner agent. In this paper, we focus on methods for maintaining continuous interaction with such agents during long-term activities. During short-term interactions such as those occurring at front desks, shopping counters, and information offices, the quality of the interaction is "mechanical," even between humans. Our aim is to develop an agent that could be regarded as a communicative social partner like human.

It is important that the mental state of people when they interact with the agents is the same as that when they interact with humans. The mental states that humans can be in with respect to an agent can be defined as physical stance, design stance and intentional stance [2]. When we take the physical stance, we pay attention to physical features such as the power of the motor, the spec of the display and so on. When we take the design stance, we expect that the agent works mechanically according to predefined rules. When we take the intentional stance, we consider that the agent has subjective thoughts and intentions. When a human interacts with another human, they usually take the intentional stance. In this case, they and their communication partner respect each other. When a human interacts with a machine, they usually take the design stance. In this case, they usually interact with the machine from a self-centered perspective because they do not consider that the machine has its own intentions. To establish social relationships between a human and an artificial agent, the agent has to induce the intentional stance.

When engaged in long-term human-human interactions, we actively demonstrate and share our own preferences, mental attitudes, and inner states to facilitate smooth interaction. Through our interactions with each other, we mutually estimate and construct a model of behavior. However, in many human-agent interactions, it is difficult for the human to estimate the agent's behavior model, because the agent and the human do not actively demonstrate and share a common ground. Consequently, the human is unable to apply a general human behavior model to the agent. To resolve this problem, previous researchers have attempted to approximate an agent's behavior model to a generalized human behavior model. However, in the course of a long-term interaction, we expect that the behavior model will be personalized along with the interaction. Therefore, this approach is not considered suitable for developing an agent that can be regarded as a communicative social partner. To achieve long-term, sustained human-agent interaction, the agent induces an active attitude within the human to demonstrate and share their common ground..

To develop our social partner agent, we propose methods for dynamically estimating emphasizing points (DEEP) based on verbal reactions, body movements, and physiological indices. These methods aim to support interactive decision-making [3][4] and to induce an intentional stance for active interaction between a human and an agent [5]. However, these previous studies have demonstrated that it is not sufficient to investigate the effects of a model's constructed inner state model on a human's attitude toward the interaction within a series of such interactions. This study aimed to investigate the effects of the agent's estimation of a human's preferences on the human's affective impressions within a series of interactions. We expected the provision of consistent estimation, using accumulated data on interactions, to induce a positive human attitude toward the agent and the interaction.

The paper is organized as follows. Section 2 briefly introduces previous works. Section 3 explains the outline of the proposed method which is partly described in previous works. Section 4 describes an experiment for comparing two types

 ISBN: 978-1-61208-440-4

of methods and then presents the results. Section 5 discusses the achievements and future work. We give our conclusion in Section 6.

II. Related work

Agents that collaboratively perform various tasks-such as subordinate support agents, when people perform tasks using their own initiative, and automated attentive agents, which automatically perform tasks in line with a human's wishes, have been proposed previously. In addition, some researchers have developed systems that can provide proposals to satisfy a user's demands. These systems gradually estimate user demand throughout the interaction.

Aydogan et al. [6] proposed an architecture in which both consumers and producers use a shared ontology to negotiate services. Through repetitive interactions, the provider accurately learns consumer needs to provide better-targeted offers. The system learns consumer needs through long-term interactions; however, it did not consider that user demands and needs could change during the process of the interaction. Azaria et al. [7] considered a two-player game in which an agent repeatedly supplies advice to a human user followed by an action taken by the user that influences both the agent's and the user's costs. That study consisted of a repeated setting that is analogous to choice-selection processes, in which a person is asked to choose a route to work from a set of possible candidates. In their study, they proposed an agent that models human behavior by combining principles from behavioral science with machine-learning techniques. In these studies, the researchers considered the effectiveness of task performance; however, they did not consider the influences on the affective attitude of the users. If a user could effectively perform a task but had poor affective impressions, the user would be hesitant to use the system or the agent. Papangelis et al. [8] argued that rapport, which is an affective attitude, had been identified as an important factor in human task.

Some studies have investigated how an agent's advice is accepted by users. Goetz et al. [9] reported that an appropriate match between a robot's social cues and its task improves people's acceptance of and cooperation with the robot. Appearance is a factor that induces affective impressions. De Melo et al. [10] explored the interpersonal effect of emotions expressed by embodied agents on human decision-making. Their results show that participants are sensitive to differences in facial displays and cooperate significantly more with a cooperative agent. These results indicate that affective impressions influence human decision-making.

III. Overview of facilitative dynamic estimation of emphasizing points with extended estimation

In an earlier study [3], we proposed DEEP method, based on verbal reactions, body movements, and physiological indices from an interaction. In Ohmoto et al. [4], we combined divergent and convergent processes with the DEEP method. We call this "facilitative DEEP" (fDEEP). In these studies, we tried to estimate "emphasizing points." The emphasizing points are factors that we consider and emphasize to reach an appropriate decision. There are many factors which influence decision-making. We implicitly focus on some of the factors and make a decision based on the focused factors. We briefly explain the method and additionally propose "extended estimation," which is needed for maintaining emphasizing points in an interactional series of tasks.

A. Estimation of emphasizing points

The degree of emphasis is rated on a scale from zero to five. The rating is changed based on the following three factors in interaction between human and a system with DEEP. The system capture the factors by using cameras, microphones, motion capture systems and a measuring system for physiological indices.

1) Verbal reactions: Either of the following two reactions occurs.

- Listed words appear in answers or demands.
- The participant provides backchanneling phrases, which express acknowledgement, surprise, or understanding, such as "ah," "oh," "aha," "I see," and "I understand."

2) Body movements: The participant repeatedly nods three times or more.

3) Physiological indices: Either of the following two responses occurs. (refer to [11], [12], [13]).

- Skin conductance response (SCR) increases more than 10% compared with resting levels.
- Low-frequency/high-frequency (LF/HF) value (electrocardiograph measurement) is more than 5.0.

Verbal reactions, body movements, and physiological indices, are used as criteria for determining when a new factor is discovered and should be emphasized, and for determining when a user's degree of emphasis of a particular factor increases or decreases.

4) Rules for changing estimated emphasizing points during interaction: A DEEP system explains the proposals and the estimated emphasizing points change depending on the participant's responses.

5) Discovery of a new factor to be emphasized: Verbal reactions, body movements, and physiological indices are the criteria for determining when a new factor is discovered and should be emphasized. When any one of the three criteria appears during interaction, the system decides that the factor should be slightly emphasized, and increases the degree of emphasis from zero to two. When any two or all three criteria are present, the system increases the emphasis from zero to three.

6) Increasing or decreasing degree of emphasis: Verbal reactions, body movements, and physiological indices are used as criteria for determining when a user's degree of emphasis of a particular factor increases or decreases. When any one of the three criteria appears, the system decides that the factor should be emphasized, and increases the emphasis of the factor by one.

When there are physiological reactions, but no verbal reactions and body movements, the system decides that the factor should be emphasized less, and decreases the emphasis of the factor by one.

7) Rules for changing estimated emphasizing points from active demands: The system asks whether or not a user has any demands. From the user's response, the system determines what the user's demands are and what changes there are to the emphasizing points. The system accepts keywords which are expected words in advance to express emphasizing points,

 ISBN: 978-1-61208-440-4

demands, and basic words necessary to capture demands in the user's responses. Words that are not expected to be included in answers are ignored.

8) Discovery of new factors to be emphasized: When the emphasis degree of the discovered factor is zero, the system increases the degree of emphasis from zero to three.

9) Increasing or decreasing degree of emphasis: When the emphasis of the discovered factor is greater than zero and the system decides that the factor should be increased, the system increases the degree by one. When the system decides that the emphasis of the factor should be decreased and the degree is greater than zero, the system decreases the degree by one.

B. Selecting the next step based on DEEP results

According to the criteria mentioned above, changes to a user's emphasizing points are estimated after the proposals are given and data are collected from the user's reactions and responses. After the estimation, the next two proposals are selected based on the estimation results.

The next proposals are selected using a table of orthogonal arrays prepared in advance. Orthogonal arrays are a special set of Latin squares, which can be used to estimate main effects using only a few experimental runs. Each proposal in the table has parameters of emphasizing points. From the table, the proposal that most satisfies the user's emphasizing points is selected. When many proposals in the table can satisfy a user's emphasizing points, a proposal is selected according to predefined rules. The rules are designed by hand. For example, the system selects a nearest proposal in convergent process because the system knows which factor is important for the user. The distances of the proposals are calculated by cosine similarity.

C. Method to control divergent and convergent processes in an interaction

The agent which supports the user's decision-making during the interaction needs to use social signals for active listening and teaming to control divergent and convergent processes based on the estimated emphasizing points in the interaction. For the facilitative interaction, we combined divergent and convergent processes with the DEEP method (fDEEP). The used signals are the frequency of providing a new proposal, recommendation from the agent, mimicry of nodding motions, and utterances.

1) The agent's behavior in the divergent process: The agent provides a small nod once in reaction to the user's utterance. The frequency of providing a new proposal is low. The agent provides a new proposal after the agent explains three emphasizing points. The furthest proposal from the previous one is selected as a new proposal. The degree of emphasis decreases if the emphasizing point is not explained in the previous proposal.

2) The agent's behavior in the convergent process: The agent provides two large nods in reaction to the user's utterance. The frequency of providing a new proposal is high. The agent provides a new proposal after the agent explains one emphasizing point, which is a recommendation. The nearest proposal to the previous one is selected as a new proposal. The degree of emphasis decreases only when the emphasizing point is clearly refused in the previous proposal.

3) The rules to switch between the divergent process and convergent process: The agent starts the interaction with a divergent process. The agent switches from the divergent process to a convergent process when the agent detects the following situations:

- There are more than three emphasizing points, with a degree of emphasis of more than one, and the degree of emphasis does not change during the interaction.
- The user offers a convergent opinion such as "I want to see like this one" and "I want to determine."

4) The emphasizing points of the agent: The agent has the same set of emphasizing points for the decision making. The emphasizing points and the degree of emphasis are the subjective opinions of the agent. The emphasizing points are set to the values of the recent proposal at the time when the agent switches from the divergent process to the convergent process. This means that the agent searches the neighbor of the last proposal of the divergent process during the convergent process. The degree of emphasis decreases when the emphasizing point is clearly refused by the user.

D. Extended estimation through maintenance of emphasizing points within a series of interactions

In this study, we used historical estimated emphasizing points within a series of tasks to estimate emphasizing points within a new but similar task. We have termed this "extended estimation," which is similar to near transfer. There are some previous studies about knowledge transfer [14][15]; however, we cannot apply the theory directly to our actual agent system. Figure 1 illustrates the concept of extended estimation. In this example, a person is coordinating their new living space. The person first selects furniture and electronics for the living space and then plans where to place them within this space. When selecting these items, their qualities are a primary consideration. During the planning phase, the person considers the relationship between the furniture and electronic items. Thus, while selection and planning are different tasks, they are correlated. Often, an adviser who helps plan the living space estimates emphasizing points during the planning phase based on the history of those involved in the selection. This individual then offers advice based on the extended estimation. One purpose of the extended estimation is to adjust the degree of each emphasizing point in the history of the previous tasks that is to be applied in the estimation of the emphasizing points of the next task. This adjustment is based on relationships that exist between these sequential tasks and it plays a role in converting the meaning of each emphasizing point within the previous and the next tasks.

To implement the extended estimation, we have added two components to DEEP. The first is for maintaining the history of the estimated emphasizing points within a series of interactions related to sequential tasks. The second adjusts the degree of each emphasizing point to a new task. We termed this enhanced version "facilitative DEEP with extended estimation" (feeDEEP). When applied to a new task, feeDEEP converts the degree of estimated emphasizing points in previous tasks to those in the new task using predefined rules, such as a predesigned relational network. For this study, we constructed predefined rules from the observations performed during preliminary experiments. In the preliminary experiment, we listed candidates of emphasizing points and the relationships among them. We interviewed the participants to select and determine

 ISBN: 978-1-61208-440-4

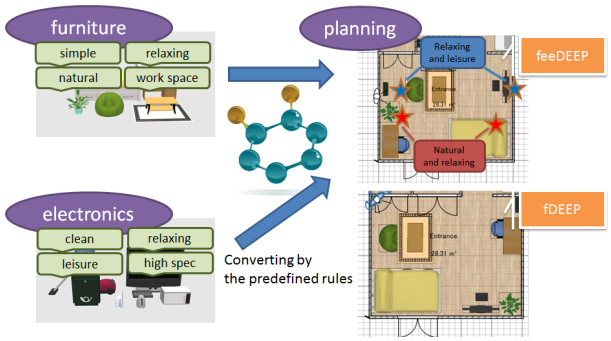


Figure 1. Example of extended estimation.

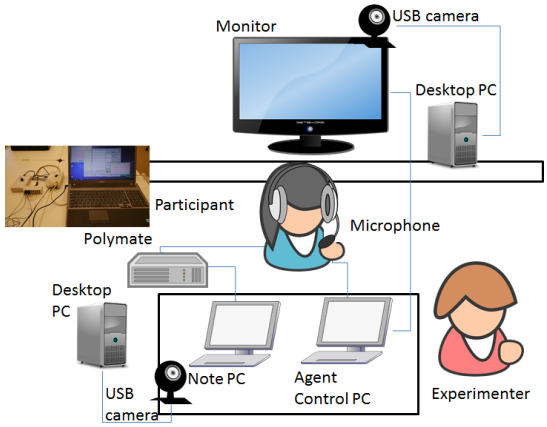


Figure 2. The experimental setting.

the emphasizing points and the weight of the relationships after the preliminary experiment. In addition, when the extended estimation achieved a degree entailing a sufficient number of emphasizing points, an agent implemented feeDEEP by initiating the interaction through a convergent process.

We expected that feeDEEP can estimate the degree of emphasizing points effectively because their initial values could be set based on user preference. The agent implementing feeDEEP could also provide affective impressions that conveyed an impression of the agent's consistent individuality to the user. These affective impressions are important for motivating users to engage in continuous interaction with the agent.

IV. Experiment

The purpose of our experiment was to investigate how feeDEEP affects the efficiency of the decision-making process and impressions related to agent behavior in an interactional series of tasks. We expected that the extended estimation will reduce cognitive load on decision-making by effective support; thus, participants would notice agent proposals that were personalized for them. We used two types of agents to evaluate the effects of the proposed method. One was an agent that implemented feeDEEP and the other was an agent that implemented fDEEP. The feeDEEP agent maintained the participant's emphasizing points; however, the fDEEP agent did not. The inputs for the agents were captured automatically, with the exception of data related to verbal meanings, such as the varbal negative feedback and the questions by the participants, because we could not robustly determine them automatically in real time, such as whether a user's utterance was positive or negative and whether the user's utterance was a question. We refer to the agent control method using manual inputs as a Wizard of Oz (WOZ) method. After the experiment, we analyzed the interaction behavior of participants and questionnaire responses.

A. Task

The participants were asked to coordinate a new living space. The primary task included three tasks, i.e., furniture selection, electronics selection, and living space planning. They first selected furniture and electronic items for the living space, and then planned where to place these items within the space. The participants interacted with the agent about the selection of items and the planning. We identified 16 factors of emphasizing points, such as relaxing, natural, for work, clean, leisure, high spec and so on, that the participants considered when they performed each task. The factors in each task were partially the same; however, some differed. In addition, between the selection tasks and the planning task, the meanings of the same factors differed.

B. Experimental setup

The experimental setting is shown in Figure 2. A participant sat in front of a 60-inch monitor displaying the agent and the proposal. The experimenter sat out of view of the participant and entered the stimuli via a WOZ interface. Two video cameras recorded participant behavior, i.e., one was placed on the monitor to record the participant's behavior and another was placed behind the participant to record the agent's behavior. The participant's voice was recorded using microphones placed under the monitor. Polymate was used to measure SCR and an electrocardiogram. The experimenter instructed the participants to keep their left arm on an armrest.

C. Procedure

After brief instruction about the experimental procedures, the experimenter showed a video of typical interaction with the agent in a preliminary experiment in which participants performed a different task. Then, electrodes were attached to the participant to measure SCR and LF/HF values. After a 2-min relaxation period, the experimenter began the first task. The participant repeatedly asked questions about the proposal and considered the proposals provided by the agent until one of the proposals satisfied the participant. The participants rested between each task. At the conclusion of the experiment, the participants completed a questionnaire.

The participants in this experiment were 21 Japanese college students (16 males and 5 females) aged between 19 and 31 years (average age was 22.9). Eleven participants (8 males and 3 females) interacted with the feeDEEP agent (feeDEEP group) and the rest interacted with the fDEEP agent (fDEEP group).

D. Results of the analysis of the number of proposals

To investigate whether the extended estimation contributed to effective decision-making, we counted the number of proposals from the agent in the second selection task and the

 ISBN: 978-1-61208-440-4

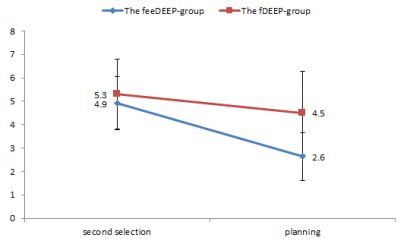


Figure 3. The number of proposals in the second selection task and in the planning task.

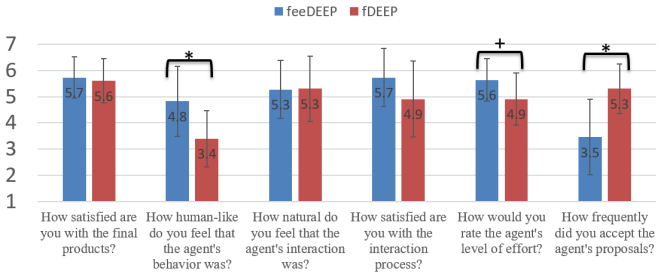


Figure 4. Questionnaire results.

planning task. We performed a paired t-test on the data of each group to compare the interaction efficiency in the tasks. Figure 3 shows a significant difference in the feeDEEP group ($p = 0.00034$); however, there was no significant difference in the fDEEP group ($p = 0.26$). In addition, we performed an unpaired t-test on the data in the planning task between each task. As a result, the value for the feeDEEP group was significantly less than that for the fDEEP group ($p = 0.0078$). These results indicate that the feeDEEP agent achieved more effective decision-making support in the planning task than the fDEEP agent.

E. Results of questionnaire analysis

The purpose of this analysis was to investigate how the extended estimation influenced the participants' subjective impressions relative to the agent's behavior. The participants answered six questions using a seven-point scale. The scale was presented as seven ticks on a black line without numbers (scored from 1 to 7). The results are shown in Figure 4. We performed a Mann-Whitney U test on the data from the questionnaire.

1) How satisfied are you with the final products?: There was no significant difference between the groups ($p = 0.53$). The average scores were higher than the midpoint of the scale; thus, the participants accepted the final products provided by the agents. This indicates that the algorithms that controlled the human-agent interaction and provided candidates were accepted by the participants.

2) How human-like do you feel that the agent's behavior was?: The participants in the feeDEEP group felt that the agent was significantly more human-like than the participants in the fDEEP group ($p = 0.035$). We believe that one reason for this is that the feeDEEP agent was consistent in terms of suggestions and candidates. Consistency by the extended estimation affected the participants' impressions of a human-like agent.

3) How natural do you feel that the agent's interaction was?: There was no significant difference between the groups ($p = 0.57$). The average scores were higher than the midpoint of the scale; thus, the agents achieved natural interaction to some degree. However, the extended estimation did not affect the participants' impressions.

4) How satisfied are you with the interaction process?: There was no significant difference between the groups ($p = 0.54$). The average scores were higher than the midpoint of the scale; thus, the participants felt relatively good about the agents' estimations. However, the extended estimation did not affect the participants' impressions.

5) How would you rate the agent's level of effort?: This question was asked because we want to know whether the participants regarded the agent as an independent-minded partner. The participants felt that the feeDEEP agent tried harder than the fDEEP agent ($p = 0.087$, marginally significant difference). Only one participant in the feeDEEP group scored lower than the midpoint of the scale, and this participant indicated that the agent's voice did not express hard work. This reason did not relate to the agent's interaction behavior. On the other hand, four participants in the fDEEP group scored higher than the midpoint of the scale. We suggest that the extended estimation affected the participants' impressions of the agent's character.

6) How frequently did you accept the agent's proposals?: The participants in the feeDEEP group accepted significantly fewer proposals than the participants in the fDEEP group ($p = 0.013$). In the interviews, most participants from the feeDEEP group said that they could express their opinions. In contrast, most of the participants in the fDEEP group said that they compromised due to the agent's ability. The ability of the agents was the same, with the exception of the extended estimation of the feeDEEP agent; therefore, we suggest that extended estimation can induce participant opinions.

V. DISCUSSION

We conducted an experiment to investigate whether extended estimation influences the efficiency of decision making and the affective impressions of the implementing agent. The experimental results indicate that extended estimation reduced the number of interactions in the decision-making process. This reduction was not attributed to the tediousness of the interaction with the agent because the extended estimation also provided positive affective impressions related to the agent's character. In addition, we did not obtain high scores for impressions related to the interaction process (naturalness and confidence). One of the reasons for this was that the agent only provided proposals for the decision-making task. Naturalness and reliability are important for motivating sustained interaction; therefore, we will consider how to gain user confidence in future research.

A particularly important finding emerged from our analysis of the data obtained from the questionnaire used in the experiment, i.e., the rate of participant acceptance of the feeDEEP agent's proposals within these interactions was significantly

 ISBN: 978-1-61208-440-4

low despite a decrease in the number of interactions. Intuitively, one of the reasons why the number of interactions may have decreased was that the agent was able to provide proposals that matched participant preference, thereby leading to quick acceptance of a proposal. However, this contradicts the results of the previous questionnaire. In addition, there was no significant difference in the number of interactions during the selection sessions. However, participants chose their top three emphasizing points out of a total of 15 at the end of the experiment. The average number of matched emphasizing points was 1.67, and all of the participants' emphasizing points had at least one match. This indicates that the agents were, to some extent, able to estimate participants' preferences accurately. Based on these results, we suggest that the feeDEEP agent induced an active attitude toward the decision-making interaction, which reflected the participants' desires due to the extended estimation. However, participants who interacted with the fDEEP agent demonstrated a passive attitude by selecting an agent's proposal that relatively matched their preferences. We further suggest that the reason for an active attitude among participants was that they were able to construct an estimation behavior model for the agent. This was possible because the agent provided a consistent estimation of emphasizing points for each participant. The participants' attitudes indicate that they regarded the agent as communicative. To create a system that is user-centric, it is necessary for the user to maintain an active attitude toward the decision-making interaction in order to accomplish their goals. This study has demonstrated one method that can be applied to induce an active attitude.

In our previous work [16][17], we analyzed physiological indices (SCR and LF/HF values) that were obtained experimentally. However, we could not include an analysis of these indices in this paper. In the near future, we will analyze these data in detail to investigate the underlying reasons for these experimental results.

VI. CONCLUSIONS

In this study, we investigated the effects of the consistent estimation of a human's preferences on the human's affective impressions within a series of human-agent interactions. For this purpose, we proposed the estimation model with extended estimation based on DEEP. We conducted an experiment to evaluate the effect of the method using two agents; a feeDEEP agent, which was proposed in this study, and a fDEEP agent, which was proposed in our previous work. The results showed that feeDEEP agent could reduce the number of interactions in the decision-making process and improved some of the affective impressions related to the agent's character. In addition, we found that the rate of participants' acceptance of the feeDEEP agent's proposals was significantly low. This was possible because the agent provided a consistent estimation of emphasizing points for each participant. The participants' attitude indicates that they regarded the agent as being communicative.

In this study, we constructed predefined rules from observations, which is one of the most important point of extended estimation. We will develop the automated construction method of the rules based on the obtained data of actual interaction in future.

ACKNOWLEDGMENT

This research is supported by Grant-in-Aid for Young Scientists (B) (KAKENHI No. 25870353), Grant-in-Aid for Scientific Research (A) (KAKENHI No. 24240023) and Grant-in-Aid for Scientific Research on Innovative Areas (KAKENHI No. 26118002) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

[1] B. Shneiderman and P. Maes, "Direct manipulation vs. interface agents," interactions, vol. 4, no. 6, 1997, pp. 42–61.

[2] D. C. Dennett, The intentional stance. MIT press, 1989.

[3] Y. Ohmoto, T. Miyake, and T. Nishida, "Dynamic estimation of emphasizing points for user satisfaction evaluations," in Proc. the 34th Annual Conference of the Cognitive Science Society, 2012, pp. 2115–2120.

[4] Y. Ohmoto, M. Kataoka, and T. Nishida, "The effect of convergent interaction using subjective opinions in the decision-making process," in Proc. the 36th Annual Conference of the Cognitive Science Society, 2014, pp. 2711–2716.

[5] Y. Ohmoto, J. Furutani, and T. Nishida, "Induction of intentional stance in human-agent interaction by presenting agent behavior of goal-oriented process using multi-modal information," in COGNITIVE 2015: The Seventh International Conference on Advanced Cognitive Technologies and Applications. IARIA, 2015, pp. 90–95.

[6] R. Aydogan and P. Yolum, "Learning consumer preferences using semanticsimilarity," in AAMAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems. New York, NY, USA: ACM, 2007, pp. 1–8.

[7] A. Azaria, Y. Gal, S. Kraus, and C. V. Goldman, "Strategic advice provision in repeated human-agent interactions," Autonomous Agents and Multi-Agent Systems, 2015, pp. 1–26.

[8] A. Papangelis, R. Zhao, and J. Cassell, "Towards a computational architecture of dyadic rapport management for virtual agents," in Intelligent Virtual Agents. Springer, 2014, pp. 320–324.

[9] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behavior to tasks to improve human-robot cooperation," in Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003. The 12th IEEE International Workshop on. IEEE, 2003, pp. 55–60.

[10] C. M. De Melo, P. Carnevale, and J. Gratch, "The influence of emotions in embodied agents on human decision-making," in Intelligent virtual agents. Springer, 2010, pp. 357–370.

[11] T. Lin, M. Omata, W. Hu, and A. Imamiya, "Do physiological data relate to traditional usability indexes?" in Proceedings of the 17th Australia conference on Computer-Human Interaction, 2005, pp. 1–10.

[12] U. M. Nater, R. La Marca, L. Florin, A. Moses, W. Langhans, and et. al., "Stress-induced changes in human salivary alpha-amylase activity?associations with adrenergic activity," Psychoneuroendocrinology, vol. 31, no. 1, 2006, pp. 49–58.

[13] K. Nakazono, T. Hada, E. Ataka, H. Tanaka, and Y. Nagashima, "Workload evaluation of gaming task by physiological indices and psychological indices," Technical report of IEICE. HIP, Tech. Rep., 2008.

[14] M. Wilkesmann and U. Wilkesmann, "Knowledge transfer as interaction between experts and novices supported by technology," Vine, vol. 41, no. 2, 2011, pp. 96–112.

[15] S. S. Gunasekaran, S. A. Mostafa, and M. S. Ahmad, "Knowledge transfer model in collective intelligence theory," in Advances in Intelligent Informatics. Springer, 2015, pp. 481–491.

[16] Y. Ohmoto, A. Matsumoto, and T. Nishida, "The effect of alternating propagation of local objective and global purpose by a network-connected two-layer model of emphasizing factors," in 2015 EuroAsian-Pacific Joint Conference on Cognitive Science, 2015, p. (in press).

[17] Y. Ohmoto, S. Takeda, and T. Nishida, "Distinction of intrinsic and extrinsic stress in an exercise game by combining multiple physiological indices," in 2015 7th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES), 2015, p. (in press).

 ISBN: 978-1-61208-440-4

Can We Monitor Crew Situational Awareness During Flight?

Exploring the use of behavioural markers

Tanja J. J. Bos
Training, simulation and operator performance
National Aerospace Laboratory NLR
Amsterdam, The Netherlands
e-mail: Tanja.Bos@nlr.nl

Antoine J. C. de Reus
Training, simulation and operator performance
National Aerospace Laboratory NLR
Amsterdam, The Netherlands
e-mail: Antoine.de.Reus@nlr.nl

Abstract— Is it possible to automatically and non-intrusively observe flight crew behaviour in the cockpit in order to monitor their Situational Awareness during flight? This work in progress investigates the possibility to automatically monitor flight crew Situational Awareness (SA). The monitoring tool is to automatically analyse the Situational Awareness of both pilots on the basis of their visual scanning, their interaction with cockpit systems and their speech. Visual scanning is an indicator for the first level of Situational Awareness (perception), which is a necessary basis for higher levels of Situational Awareness (comprehension and projection). The timing of pilot interactions with cockpit systems as well as speech could be indicators of these higher levels of Situational Awareness. The question we would like to answer in our work is: can we establish a common behavioural pattern within a flight crew that indicates optimal SA and can we use this as a reference to identify reduced SA? And, if this proves to be impossible, is an individual reference pattern possible and is that a suitable alternative?

Keywords- crew monitoring; Situational Awareness; modelling.

I. INTRODUCTION

Continuous combined efforts by the aviation community have resulted in the achievement of a safety record in air transport that is unequalled by other modes of transport. In recent years, although these efforts continue unabated, the accident rate seems to have reached a stable rate of about 2 accidents per ten million flights [1]. With air transport foreseen to grow in the coming decades (4.7% to 5.0% per year), this accident rate, while very low, will translate to several major incidents and accidents per week.

A fundamental pillar in aviation safety is the requirement that no single failure should result in a catastrophic accident, i.e., the write-off of an aircraft and multiple fatalities. During aircraft type certification the manufacturer must prove to the aviation authorities that the probability of a system failure is below a threshold that is inversely related to the severity of the consequences. For example, potentially catastrophic failures must have a probability of less than one occurrence in every 100.000.000 flight hours.

Aircraft systems have become so reliable, that human error has become more prominent in statistical analyses of factors contributing to accidents. Thus, a large percentage of recent accidents can be linked somehow to a human factors issue, such as poor perception of the environment, inadequate crew coordination, excessive workload, misunderstanding of an evolving situation and inappropriate training.

Even in case of two well-trained pilots, complex and high workload conditions are not uncommon today. Especially in these conditions, optimum crew action depends on an adequate understanding of the situation at hand and the corrective actions that are needed. Monitoring crew status in flight is therefore considered one of the enablers for enhancing overall safety. When suboptimal crew status is indicated, the crew could be assisted. In this work, we particularly look at crew Situational Awareness (SA).

The overall objective of the work is to automatically monitor the pilots' state in a non-intrusive way. This includes constructs such as flight crew's presence in their seats, physical state, drowsiness, workload, distraction, etc. Situational Awareness is one of these constructs. So far, it is not possible to directly measure Situational Awareness in a non-intrusive way. Thus, Situational Awareness is usually assessed using ratings by an observer or using self-ratings and questionnaires, such as Situation Awareness Rating Technique (SART) and Crew Awareness Rating Scale (CARS). These measurements are neither automatic nor non-intrusive and not suitable for day to day use in an airline cockpit. In this operational environment, Situational Awareness should be inferred from automatic behavioural analysis of non-intrusive measurements.

This paper describes the work in progress. The sections included describe: the concept of SA (2), the operational context (3), the potential observables for SA (4), the implementation of the tool in the operational context (5), the method used (6) and the last section the status at the time of writing.

 ISBN: 978-1-61208-440-4

II. SITUATIONAL AWARENESS

One of the most widely accepted definitions of Situational Awareness is that of Endsley [2]: *"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"*.

This definition includes three levels of SA:

- *Level 1, Perception* – This is the basic information that is required for Situational Awareness, being able to notice events, people, objects in the external environment. This level simply represents the collection of basic data.
- *Level 2, Comprehension* – Once an event, or object, has been perceived, it is necessary to understand the meaning of that object in the situation. This represents the interpretation of the basic data that is collected through perception.
- *Level 3, Projection* – The third level of Situational Awareness represents the ability to project the interpretation of the current situation into the future. This level of Situational Awareness is required to be able to predict the effect of the information that is currently available onto the future situation.

Situational Awareness of pilots has been assessed during many flight simulation experiments, but monitoring it automatically and non-intrusively in an operational environment is the challenge of this work in progress.

III. THE PROCEDURAL CHARACTER OF FLYING

The tasks of a pilot consist of actions necessary to fly the aircraft, to navigate, to communicate and to manage systems. Crew behaviour in the cockpit is largely driven by procedures. These procedures are dependent on the flight phase and are influenced by the environment (e.g., weather, terrain), external events (e.g., ATC commands) and the aircraft state (e.g., speed, fuel, systems status).

Crew behaviour is also driven by crew resource management. One pilot is responsible for flying the aircraft and the other for other tasks such as the communication (with the cabin and outside world), and managing systems. The pilots in their roles have complementary tasks, but both need to assure themselves that primary flight parameters are within the acceptable range. Together they are responsible for a safe flight execution.

Descents, approaches and landings are the more busy flight phases, which consist of a relatively predictable number of actions and the use of checklists. Depending on the type of technology available on the airport and other conditions such as visibility, the type of landing, the level of automation is selected. The different landings may require different procedures but it is expected that roughly the visual information acquisition behaviour is similar.

IV. OBSERVABLES

Much of the information that is available in the cockpit is of the visual modality. Hence, the crew's scanning pattern and visual focus points are an important method of observing the crew's attention. The crew's scanning pattern can be recorded with an eye-tracker. The crew's interaction with cockpit systems and speech can also be recorded directly. How suitable are these three types of behavioural measurements for automatic recording and interpretation on the flightdeck in actual flight?

A. Eye gaze

Eye gaze is highly selective in the sense that it provides direct insight in the information that is being used by pilots [3][4] for performing their tasks. Eye gaze is also highly generalizable, i.e., all pilots must visually scan cockpit instruments and the outside world in order to retrieve most of the required information. Eye gaze is highly responsive too: a single eye fixation related to information intake can be in the order of 70 ms. Finally, with state-of-the-art algorithms, eye gaze data can be reliably filtered and interpreted. Note that some displays in the cockpit can present different pages with information. Consequently, the active page should be taken into account when interpreting the data.

B. Interaction

Interaction with cockpit systems is highly selective in the sense that it provides direct insight in the tasks pilots are performing. Besides, every crew has to perform the same interactions to obtain the same results, so measures of interactions are also high generalizable. They are highly responsive too: every interaction is directly and instantly related to a pilot task. For this reason, complex filtering is not needed and automatic interpretation is straightforward. Note however, that in less busy flight phases the crew has some freedom in the order of performing their tasks.

C. Speech

Speech recognition can provide insight in the topics the pilots are discussing; however, it is more difficult to reliably infer the precise meaning of the vocalizations in relation to task performance. Note that this may change in the near future since speech recognition technology is developing rapidly (consider the "digital assistants" on smartphones, that rely heavily on understanding natural speech). Also note that interpretation of intonation can also give clues regarding crew state and this has been shown to work. However, it provides little concrete information regarding task occupation.

All in all, recordings of eye gaze and interaction with cockpit systems are currently the most suitable behavioural

 ISBN: 978-1-61208-440-4

measures for our purpose. Finally, note that these measures primarily relate to SA Level 1 and 2.

V. METHOD

As a first step in our work, the SA assessment module is scoped around one flight phase, for which the required behaviour is particularly procedural and therefore relatively predictable: the full descent under nominal conditions.
The input for the SA module will be the visual acquisition on the basis of eye-tracker information, the altitude in relation to the runway, the information that is displayed on the cockpit systems and the communication with Air Traffic Control.

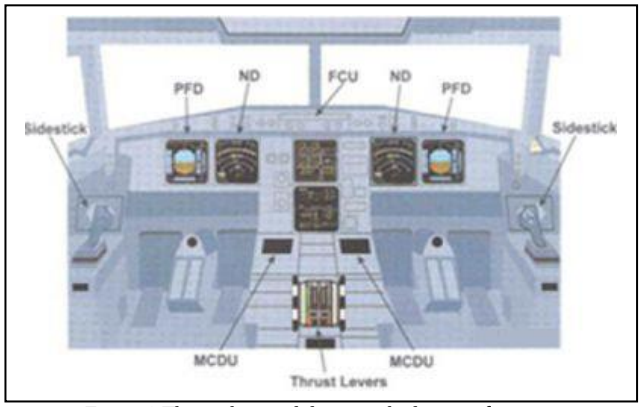


Figure 1 The cockpit and distinguished areas of interest

The methodology consists of two phases: a development phase to create the reference model and a validation [4][5] phase to validate the SA module that compares actual behaviour to the reference behaviour in real time.

A. Development phase

For the development of the reference model, pilots made descents in NLR's Airbus A320 lookalike cockpit mock-up using an eye-tracker system to register their information acquisition.

Scenarios were prepared with full descents to Schiphol airport, Amsterdam. The scenarios started a few minutes before top of descent to allow the pilots to prepare the descent and to build their SA as they would in a normal operation. Approach and landing checklists were part of the procedure. Each pilot, after a familiarisation session, flew three scenarios. One scenario concerned nominal conditions including heading instructions around a weather cell. In the other two scenarios the conditions were less optimal. In one scenario, the fuel on board was low and the landing condition as proposed by Air Traffic Control was less optimal. In the third scenarios a flap malfunction forced the pilot to divert to a different destination airport. It was observed if the pilot perceived the condition, if he understood it and took appropriate action timely.

Participants' behaviour was observed and after each scenario the pilots were interviewed to assess how they experienced the scenario and how they rated their SA in the course of the run. These observations and ratings are used in the development phase to tune the system.

The eye-tracker system delivers visual information acquisition in terms of dwells: the uninterrupted amount of time spent on an Area of Interest. The relevant Areas of interest are visualized in Figure 1. The data was analysed post experiment, in relation to the altitude above the landing runway and compared to the SA ratings and observations. The analysis allowed defining a reference model that represents adequate SA and thresholds for degraded SA. This common reference model was then integrated into the SA module to real-time monitor SA.

B. Validation phase

A group of pilots will participate to a validation exercise. In this exercise the SA module will provide indications of the pilot's SA in real time. This will also take place in a simulator that resembles an A320 cockpit.

VI. CONCLUSION AND FUTURE WORK

In attempts to further increase flight safety, tools are being developed to monitor pilots' status in flight. This work-in-progress aims to develop a module that can monitor the pilots' Situational Awareness in a non-intrusive way. For this, the pilots' visual acquisition and communication with Air Traffic Control are the observables used. A model was developed of minimal desired division of attention on sub-tasks: Aviate, Navigate, Communicate and Manage Systems during descent and landing. The model was incorporated into a module able to monitor the pilots' division of attention, and indicating degraded SA.

At the time of finalising this work-in-progress paper, the module is being evaluated in a validation exercise. This will allow further improvement of the model.

ACKNOWLEDGMENT

This work is part of the EU 7th framework project ACROSS (Advanced Cockpit for Reduction of Stress and Workload).

REFERENCES

[1] Airbus S.A.S., Quality and Safety First, Commercial Aviation Accidents, A Statistical Analysis, May 2015. http://www.airbus.com/company/aircraft-manufacture/quality-and-safety-first/ Last retrieved September 2015.

[2] M. Endsley, "Toward a theory of Situational Awareness in dynamic systems", Human Factors 37, pp. 32–64. 1995.

[3] H. van Dijk, G. K. van de Merwe, and G. D. R. Zon, Situational Awareness Assessment in Flight Simulator Experiment. Proceedings ISAP 2009.

 ISBN: 978-1-61208-440-4

[4] G. D. R. Zon, H. van Dijk, A Selection of Human Factors Tools: Measuring HCI Aspects of Flight Deck Technologies. Proceedings HCI International 2009.

[5] G. D. R. Zon, and M. I. Roerdink, HCI Testing in Flight Simulator: Set Up and Crew Briefing Procedures - Unique Design and Test Cycle. Proceedings HCI International 2007

 ISBN: 978-1-61208-440-4

Spectrum Monitoring. An Approach Based on People-Centric Sensing (SMOPEC)

Cuevas-Ruíz, J.L.
Centro de Estudios.
Instituto Federal de Telecomunicaciones. IFT.
México D.F., México
e-mail: jose.cuevas@ift.org.mx

García-Alba- Iduñate, P.
Centro de Estudios.
Instituto Federal de Telecomunicaciones. IFT.
México D.F., México
e-mail: pascual.garcía@ift.org.mx

***Abstract*— Spectrum monitoring is a key factor in the spectrum management strategy implemented by government entities. It is of enormous importance to evaluate the spectrum use and its overall performance. Spectrum monitoring, with inspection and law enforcement provide useful tools to maintain the integrity of the spectrum management process. The infrastructure and equipment necessary to get a complete monitoring system require a huge investment, about money, time and human resources. Using People-Centric sensing it is possible to develop a monitoring system using the users' mobile devices. An app installed in the mobile device allows to manage all the information obtained by the device sensors. The monitoring system based on mobile devices, would reduce the mentioned investment. Also, the system updates could be cheaper and easier. In this paper we present a Spectrum Monitoring approach using a People-Centric sensing (SMOPEC) approach; the system elements are described, the operation are detailed, a pilot test results are showed and the feasibility are exposed.**

Keywords*— *Spectrum Monitoring; People-Centric; Spectrum Management.

I. INTRODUCTION

Actually, technological advances in electronics, computation and telecommunications will turn the near ubiquitous smart phones into a global mobile sensing device, integrating an extensive sensors network for all the base station coverage area. These devices are able to measure some parameters like temperature, acceleration or user localization; also the smart phones can obtain other technical data, like reception power, transmission frequency, Signal to Noise ratio (SNR), data rate, etc. In Figure 1 some sensors and technical information obtained from smart phones are identified.

The cell network could be seen as a people-centric system, where the users' devices become the focal point of sensing and the sensor-based information is used for get a more elaborate analysis [1]. People using their mobile phones integrate sensing networks called people-centric sensing networks that sense what we are doing and support our daily activities. Most of the sensing applications depend on the ability to monitor statistics including max, average, ranking, rather than raw sensor readings [2]. These sensor networks are proposed to be used to spectrum monitoring.

Otherwise, and as a consequence of increased demand for mobile services around the world, the International Mobile Technology (IMT) spectrum has been defined as a very precious commodity. The demand for more spectrum is increasing every day. This higher demand could generate some spectrum scarcity in these bands; however, some authors mentioned that this could be just an "apparent scarcity"; they stablish that this scarcity is due to inefficient spectrum management strategies [4].

To fulfill this increased demand the government entities and operators have the commitment to integrate an efficient spectrum management strategy in order to optimize the spectrum use; this spectrum management strategy must be based on spectrum monitoring system. The monitoring system will provide an accurate and actualized information about the spectrum use. The provided information will be fundamental to integrate an efficient spectrum management strategy [3].

The information provided by the spectrum monitoring system contain the assigned bands, the interference levels, the coverage area, quality of service, etc. All these parameters will be useful for the management entities of the spectrum radio, including governments and operators. Actually, to get these information, they spent a lot of resources integrating monitoring spectrum systems, including investment in measurement equipment, human resources and time.

In this paper we present a Spectrum Monitoring approach based on a People Centric sensing (SMOPEC) system, to inventorying the spectrum utilization while this is occupied by mobile users.

This paper is organized as follows: In Section II spectrum monitoring concept are described, detailing its importance as key issue for spectrum management, in order to get more spectrum efficiency. In Section III the SMOPEC system is detailed, identifying its elements and basic operation is explained. The monitoring process based on the SMOPEC system is explained in Section IV by describing some technical issues. Some results, conclusions and future works are mentioned in Sections V and VI.

II. SPECTRUM MONITORING AND MANAGEMENT

Spectrum management is the combination of administrative, scientific and technical procedures necessary to ensure the efficient operation of radio communication equipment and services without causing interference. Simply stated, spectrum management is the overall process of regulating and administering use of the radio frequency spectrum [5]. The goal of spectrum management is to maximize spectrum efficiency, minimize interference and eliminate unauthorized

 ISBN: 978-1-61208-440-4

Figure 1. Mobile sensor network using the mobile phones. The SMOPEC concept.

and improper use of the spectrum. Rules and regulations, based on relevant legislation, form a regulatory and legal databases facilitates the spectrum management process resulting in decisions for spectrum allocations, frequency assignments, and licensing. Spectrum monitoring, inspection, and law enforcement provide the necessary means to maintain the integrity of the spectrum management process.

For government managers and mobile operators the spectrum monitoring is a key factor in order to achieve a more efficient use of this commodity. The employment of an infrastructure and organization to monitoring the spectrum operational issues could be a very expensive assignment.

For governments, the assigned frequency bands to the operators for IMT services are based on national and/or international directives; in these advices are defined technical requirements regarding to coverage area, interference level, power transmission, etc. Moreover, these directions include fines and others punitive actions; these recommendations should be applied when the operator has an omission / failure in the operation over the assigned frequency band. However, if there is not a mechanism to monitoring the spectrum performance, the operator assessment could be a so problematic mission.

Otherwise, for mobile operators, the spectrum monitoring allow them identify opportunities and “holes” in the used bandwidth (inefficient use); with this information the operator can implement new shared spectrum algorithm to increase spectrum efficiency, evaluate the Quality of Service (QoS) delivered or rate provided to users; information about coverage area and power reception could be useful to integrate plans about marketing strategies too.

For this project, the key sensors are the accelerometer (to detect user movements), Global Position System (GPS) and the information obtained from the RF signal captured by the mobile. In this first stage, we focus to get a first approach to evaluate the SMOPEC feasibility to monitoring spectrum use.

According to the above mentioned facts, the spectrum monitoring is a valuable tool in order to increase the spectrum efficiency and benefits. Monitoring is closely associated with inspection and compliance in that it enable the identification and measurement of interference sources, the verification of proper technical and operational characteristics of radiated signals, and detection and identification of illegal transmitters. Monitoring further supports the overall spectrum management effort by providing general measurement of channel and band usage, including channel availability statistics and the effectiveness of spectrum management procedures. It obtains statistical information of a technical and operational nature on spectrum occupancy. Monitoring is also useful for planning, in that it can assist spectrum managers in understanding the

 ISBN: 978-1-61208-440-4

level of spectrum use as compared with the assignments signed [7].

As the demand for mobile communication systems grows, the needs for radio spectrum increases too. In order to generate an effective and efficient spectrum management, it is imperative to possess a full understanding about how the IMT spectrum behaves over time, frequency and space. Then, the spectrum monitoring issue is a key factor in order to achieve an efficient spectrum use.

III. SMOPEC. GENERAL CONCEPT

The spectrum monitoring by people-centric strategies using the mobile devices provide a less expensive alternative to obtain the spectrum parameters performance, inside the huge investment required for the traditional approach. Moreover, due that the monitoring operation using the mobile phones are based on an app installed on the mobile device, the system is flexible, allowing that new characteristics could be added just with a software update.

The smart-phones around the world are increasing every day. Using these devices the users can receive a lot of information from practically anywhere around the world. At the beginning, the users were defined primarily just as information receivers, however this condition have been changing due to the increased connectivity available in many places and a better quality in these links. Moreover, in the smart phones are integrated a numerous sensors that can obtain a lot of the environmental information, also operational and technical data. The information collected for the sensors could be saved, processed and sent by the device to other specific user or shared in some social network. Some of these information could be sent without the user intervention, in an automatic mode. Mobile phones sending data automatically from their location integrate a sensor mobile network, and this scenario open a lot of possibilities. This is the main concept behind the SMOPEC system: thousands of mobile sensors collecting and providing information about the IMT spectrum performance. The mobile devices situated in the base station coverage area, are generating information about the spectrum performance; these spectrum information are sending to a central data processing where a more specific analysis is done.

In order to the mobile phone send just the key information, an app is installed on the device. This app manage the information obtained by the sensor, saving the data sensed, pre-processing it and sending the relevant data. The information sent from each user are: reception power (dB), location, SNR, reception frequency, standard used, call duration and data rate (bps). The pre-process and analysis are explained in the next sections. The data protection issue is not analyzed in this paper.

IV. MONITORING USING SMOPEC

As mentioned before, the installed app on the smart phone manages the collected information. These data are collected and sent by the cell phone throw the cell network to the data center, where all the information are processed. A system overview is showed in the Figure 2. The system operation is as follows:

a) Using the installed app, the device save the parameters values obtained by the cell phone sensors.
b) To avoid send redundant information to the data center processing, an algorithm compute the data correlation to select the data to be sent. This algorithm economize the operations volume performed in the device (saving energy and time).
c) The selected information is sent throw the cell net toward the processing data center; the data are analyzed and processed, using some statistical algorithms. Localization, data rate, call duration, operation frequency and power received were reported.
d) An information report is integrated and sent to the operator, indicating coverage area, average rate, users, technology, etc.

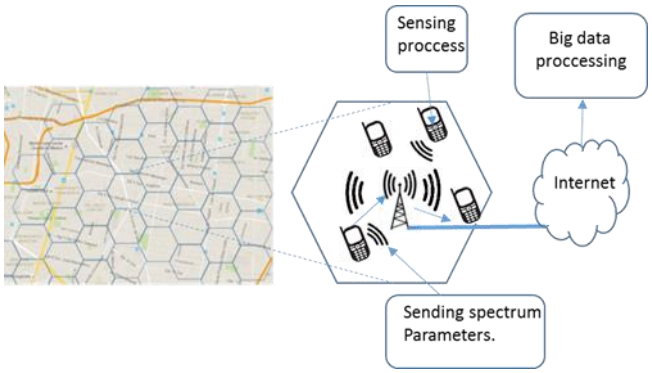


Figure 2. Spectrum Monitoring System

In the monitoring process, there are some particulars concerns to be considered to avoid that the monitoring process interfere with the device overall performance. Among others, we consider:

- Store capacity.
- Energy consume.
- Bandwidth used.

TABLE I RESULTS FROM MONITORING PROCESS

Users	Location (Lat, Long)	Power Rx (dBW)	Rate (mbps)	Call duration (secs).
1	19.398936, -99.171843	-82.7	6.2	4.27
2	19.583247, -99.236731	-75.89	1.3	1.8
3	19.266798, -99.266567	-92.25	0.8	4.9
4	19.599682, -99.227647	-88.9	5.7	8.96
5	19.398517, -99.156978	-72.49	9.2	7.02
6	19.520446, -96.92688	-89.92	3.7	2.95
7	19.479021, -96.871948	-94.26	2.9	3.76

The monitoring process is integrated for 5 stages: sensing data, pre-processing, save selected information, send the saved information and erase the sent data.

To avoid affectations on the device store capacity, the app installed in the mobile phone will send the selected data just a few seconds after the data were saved; once the information be sent to the data center processing, this information will be eliminated from the device memory, to achieve a more efficient memory use.

The operations committed with the spectrum monitoring assignment will consume extra energy; the goal is to get that this extra consumption be minimum. To get this reduction in the energy consumption, an algorithm is implemented to compute the data correlation between two consecutive samples obtained by each sensor, named as x(t) and x(t-1). In the Figure 3 is showed the flow diagram for this algorithm.

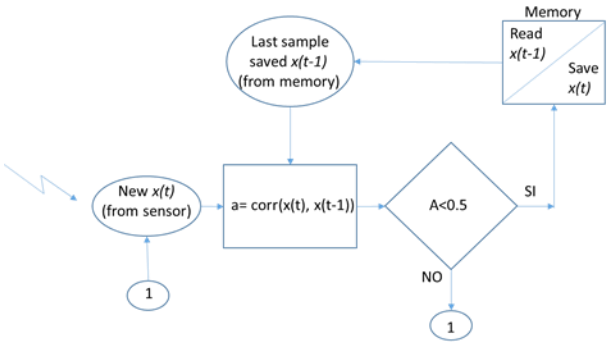


Figure 3. Monitoring process algorithm

In the monitoring process, and to optimize the memory occupancy is needed to determine if the obtained information, x(t), is different compared with the previous saved saved, x(t-1) (to avoid redundancy). To implement this evaluation, the correlation coefficient between these samples is calculated. If the correlation coefficient between the two samples is bigger than 0.5, then the system assume that the conditions in which the device are operating are the same and there are not changes about spectrum operation (i.e. same rate or same location); in this case, not new information is generated and just an update procedure is execute in the data center processing using the last information available (last sample saved). Otherwise, if the correlation coefficient is less than 0.5, then the sensed data in x(t) is defined as different compared with the previous sample saved in x(t-1), then the value in x(t) will be considered as new information as a consequence of a some changes in the operation conditions; therefore, x(t) will be saved and sent, in order to actualize the data base in the data center processing. Once the data was sent, it will be eliminated from memory as be stated before.

Considering the before explanation, just new information will be collocated over the channel. With these scheme, there is a considerable reduction in the information amount sent by the wireless channel, achieving an improved in the spectrum efficiency. In the case for GPS sensor, the app read the accelerometer values, taking a sample each 5 seconds to determine if the user is in movement or not. If there is some acceleration detected by the sensor, then it assume that the user is moving and a new location will be detected.

One advantage of the spectrum monitoring using the app installed on the mobile device is that all the parameters defined to perform the monitoring process can be changed to obtain more or less sample rate for the data collected, in order to find the ideal values to get accurate measurements.

To get a higher coverage area levels (more user using the app), the users must to obtain some benefits using it. Due that the system would provide valuable information for operators and Government, the business model must be contemplate an investment to offer some benefits to users. These paybacks could be free calls, others apps, new devices, etc. This is a key issue for the project.

 ISBN: 978-1-61208-440-4

V. RESULTS

For the pilot test, the app was installed in several smart phones and the users were monitoring while they were using their devices (voice or data). The data collected were: geographical location, power reception, data rate, and call duration. In Table 1 some results are showed. The mobile phones transmission were done using the GSM and LTE bands (800 and 1700 MHz), reporting an average rate of 3.3 Mbps. The information showed correspond a sample taken from the overall users registered in the data base; it means that the information was pre-processed and just the relevant data were sent and saved in the data base. The users in the test were distributed mainly in the metropolitan Mexican area (DF and Mexico State) and the Veracruz state. In the first case, there is metropolitan area, with a high user per area density; otherwise, for the second scenario we consider a low user density. The devices were identified by their IP (Internet Protocol) and their Mobile Identification Number (MIN). The test were running for about one week with a user average number of 150 users randomly distributed. The mobile users did not report any malfunction or some affectation over the device performance.

With collected information by the GPS system, a report with coverage area are generated; in the Figure 5 are showed two users position in one minute interval, using the available samples in the data center processing.

The final report can integrate the users per area and user per base station, among others data.

The power reception is an average power over the time that the mobile is receiving information. The power reception and the data rate are important issues to evaluate the quality of service (QoS) offered.

The sensors quality and the environmental interference could affect the information accuracy. To get a system error margin, the data collected and processed by the SMOPEC system, could be compared with the information obtained from past measurement campaigns stored in data bases.

VI. CONCLUSIONS AND FUTURE WORK

The SMOPEC project offer several opportunities. Regarding to the SMOPEC implementation, the software that manage the collected information can be installed with the operative system from factory. In this way, all the mobile phones will be able to be added to the mobile sensors network and consequently, more information will be obtained.

For governments, this alternative brings some opportunities. An application can be developed to monitor the frequency bands dedicated to official and public services. Moreover, the application will send only the parameters that the authority decides. Thus, to prevent the key and sensitive information was exposed.

Big Data algorithms can be used to process all the information and new opportunities can be developed. For example, in marketing areas, the user behavior and consumer habits would be predicted.

In some cases, some interference signals were detected. In order to determine if these transmissions were done by a license transmitter, the information could be saved and compared with other interference signals captured before in the same geographical area. When these signals were highly correlated, the system could report as a probable illegal transmitter. The authority might compare the signal reported (magnitude and time), with the legal communications done in that area to determine if it was a license transmission or not.

REFERENCES

[1] A. T. Campbel, S. B. Eisenman, D. D. Lane, E. Miluzzo and R. A. Peterson. "The Rise of People-Centric Sensing". Internet Computing, IEEE. Year: 2008, Volume: 12, Issue: 4.

[2] Motegi Shinji, Nishimra Yasutaka and Yoshihara Kiyohito. "Communication Algorithm for Statistics Monitoring in People-Centric Sensing Networks". 2010. Networks (ICN), 2010 Ninth International Conference on.

[3] A. T. Campbel, S. B. Eisenman, D. D. Lane, E. Miluzzo and R. A. Peterson. "The Rise of People-Centric Sensing". Internet Computing, IEEE. Year: 2008, Volume: 12, Issue: 4

[4] Rep. ITU-R SM. 2012-1. "Economic aspects of spectrum management".

[5] A. Chandra. "Spectrum management for future generation wireless based technology". Wireless Technology Conference, 2009. EuWIT 2009. European Year: 2009 pp. 201 – 205.

[6] OECD Digital Economy Pappers. "New Approaches to Spectrum Management". No. 235, OECD Publishing.

[7] S. W. Boyd, J. M. Frye, M. B. Pursley and T. C. Royster. "Receiver Statistics for Spectrum Monitoring While Communicating". Global Telecommunications Conference, 2009. IEEE GLOBECOM. .

 ISBN: 978-1-61208-440-4

Development of a Real-Time Assist System Increasing Walking Efficiency

Shotaro Murakami
Intelligent Mechanical Systems Engineering Course
Kochi University of Technology
Kami, Japan
e-mail: 185055r@ugs.kochi-tech.ac.jp

Yoshio Inoue
Research Institute of KUT
Kochi University of Technology
Kami, Japan
e-mail: Inoue.yoshio@kochi-tech.ac.jp

Kyoko Shibata
School of systems engineering
Kochi University of Technology
Kami, Japan
e-mail: shibata.kyoko@kochi-tech.ac.jp

Kunihiko Tachibana
School of systems engineering
Kochi University of Technology
Kami, Japan
e-mail: tachibana.kunihiko@kochi-tech.ac.jp

Abstract— **The aim of this study is to develop a gait assist system to lead users to energy-efficient walking. Energy consumption of walking is estimated by measuring both step length and stride frequency. Thus, in this paper, a test model to assist users based on measured step length and stride frequency in real-time is developed. Walking experiment is carried out and the validity of the proposed measurement method is confirmed.**

Keywords-energy comsumption; step length; acceleration sensor; gyroscope; potentiometer .

I. INTRODUCTION

Walking is the most important means of migration and is subjected to study from many different perspectives for years. In particular, a large number of articles on energy consumption of walking are published in the world, some of which are included in [1] to [3]. In addition, Farris et al. [4] and Kobetic et al. [5] explain that walking assist devices are being developed. Most of these devices are still in the research phase, but some of them are in practical use; for example [6] introduces "Walking Assist Device with Stride Management system", which is an assist device to reduce energy consumption. However, there still remains a difficult problem and that is the limitation on the gait of the user in calculating the step length only by hip joint angle.

In this paper, we present an assist system using an inertial sensor in order to develop an assist device that can be adapted to a wide range of the gaits.

In Sections 2 and 3, we explain the most appropriate relationship between step length and stride frequency, and present the method to measure them. After that, we present the test model that measures step length and stride frequency and could assist users in real-time in Section 4. Finally, in Section 5, we report the results of the verification of accuracy calculation method and measurement by test model.

II. ENERGY CONSUMPTION OF WALKING

Diedrich and Warren [1] and Molen et al. [2] show in Figure 1 that the energy consumption of walking is estimated by measuring step length and stride frequency. The most appropriate relationship between step length and stride frequency is computed from the following equation:

$$s = 0.0063n \quad (1)$$

where s (m) is step length and n (steps/min) is stride frequency.

Measurement tests were conducted using expiration gas analyser (PG-240) in order to determine the validity of the findings of [1] and [2]. Energy consumption of walking is measured changing stride frequency in different conditions: 1. The subject walks at 3.0 (km/h), 2. The subject walks at 6.0 (km/h). In addition, gait speed and stride frequency were determined by a treadmill and a metronome.

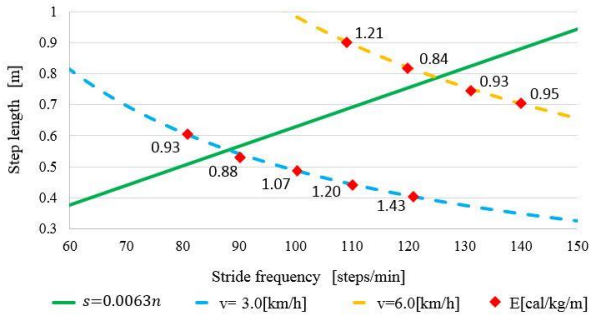


Figure 1. Verification of accuracy of the findings of [1] and [2]

In Figure 1, the green solid line indicates the most appropriate relationship between step length and stride frequency, and the red diamonds indicate energy consumption per unit distance (cal/kg/m). We can also see that the energy consumption of walking decrease when the relationship between step length and stride frequency is close to the green solid line.

 ISBN: 978-1-61208-440-4

III. MEASUREMENT OF STEP LENGTH AND STRIDE FREQUENCY

In this study, step length and stride frequency are determined by joint angles while walking, and joint angles are measured by potentiometers and encoders (Figure 2). The model of lower limb and variables in calculation are shown in Figure 2.

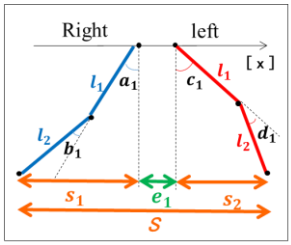


Figure 2. Variables in calculating step length

Walking cycle T(s) is easily determined because walking is a cyclic movement. Therefore, stride frequency is computed from the following equation:

$$f = 1/T \tag{2}$$

Right stride is computed from the following equation:

$$s_1 = l_1 \sin a_1 + l_2 \sin(a_1 + b_1) \tag{3}$$

where l_1 (m) is thigh length, l_2 (m) is lower length, a_1 (rad) is hip joint angle, and b_1 (rad) is knee joint angle. In addition, left stride is computed in a similar way. Then, an effect of twisting of the lower back is added. Finally, the step length is computed from the following equation:

$$S = s_1 + s_2 + e_1 \tag{4}$$

In addition, foot contact, which decides the timing of the calculation, stride frequency and an effect of twisting of the lower back can be determined by an inertial sensor on the lower back.

IV. MECHANISM OF TEST MODEL

Figure 3 is the test model that measures step length and stride frequency and could assist users in real-time.

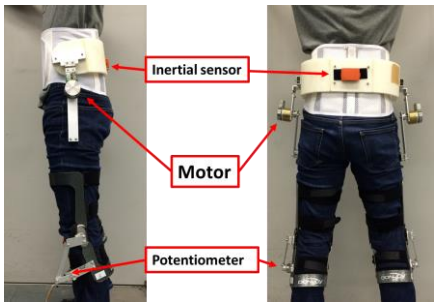


Figure 3. Test model

Motors with encoders were positioned on the axis of the hip joints for the purpose of measuring hip joint angles. Two potentiometers were positioned on the axis of the ankle joints for the purpose of measuring hip joint angles. An inertial sensor that has an acceleration sensor and a gyro sensor were positioned on lower back for purpose of measuring a longitudinal-direction acceleration value and an effect of twisting of the lower back.

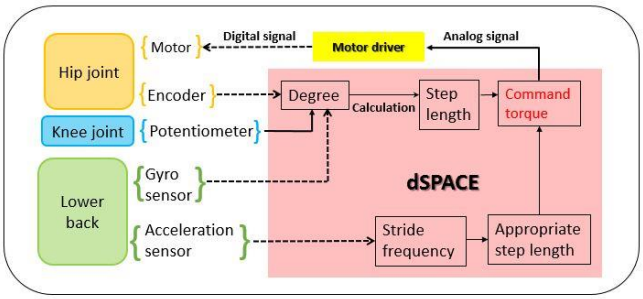


Figure 4. Real-time control system using dSPACE

Figure 4 shows the real-time control system. A PC is loaded with MATLAB/Simulink, a dSPACE data acquisition and real-time control system. The controller is designed in MATLAB/Simulink. Command torque is given to motor driver as analog signal. This system could measure step length and stride frequency and assist users in real-time.

V. VERIFICATION OF ACCURACY OF CALCULATION METHOD AND MEASUREMENT BY TEST MODEL

In order to confirm the validity of the proposed method in section 3, the accuracy of measurement was verified by using 3D motion analysis (Hawk Digital Real-Time System：Motion analysis). The results of the calculations and 3D motion analysis are compared.

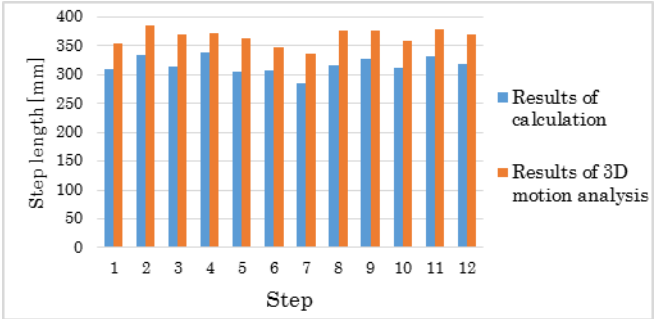


Figure 5. Step length

Figure 5 shows that the measurement error is 2.57(%) and standard deviation is 1.54(mm). The method that proposed had limitation of measurement accuracy. However, the results of calculation could be corrected by using the decided correction coefficient because standard deviation is very small.

Measurement tests were conducted using the test model. First, the floor was marked at 1 meter intervals. Second, the subject walked using regular steps using the markers as a guide. The subject was asked to walk 5 meters in 5 steps. Third, the result of calculation was compared to the actual walking distance.

 ISBN: 978-1-61208-440-4

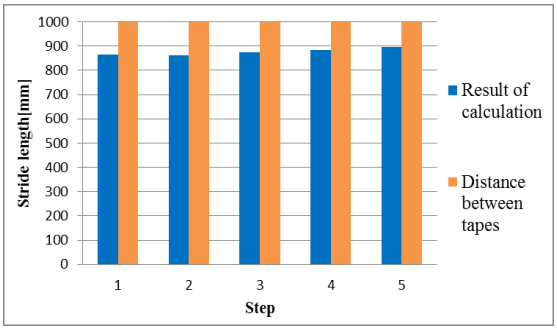


Figure 6. Step length measurements by test model and tape pith

Figure 6 shows that the measurement error is 12.4(%) and standard deviation is 12.6(mm). A main error source may be the walking movement misalignment between the potentiometers and legs. However, this result is positive because improvement of accuracy is expected by multiplying calculation result by calibration coefficient.

VI. CONCLUSION

In conclusion, a method for measuring step length and stride frequency was proposed and the test model that could measure step length and stride frequency, and assist users in real-time was developed. In addition, verification of measurement accuracy was conducted using the test model. The results show that the proposed method to determine walking efficiency was reasonable. In order to find out the optimal assist timing and the size of motors torque of this assist method, future experiments with the test model need to be conducted.

REFERENCES

[1] F. J. Diedrich, and W. H. Jr. Warren, "Why Change Gaits? Dynamics of the Walk-Run Transition," Journal of Experimental Psychology: Human Perception and Performance, Feb. 1995, Vol. 21, No. 1, pp. 183-202.

[2] N. H. Molen, R. H. Rozendal, and, W. Boon, "Graphic representation of the relationship between oxygen-consumption and characteristics of normal gait of the human male," Proceedings Koninklijke Nederlandse Academie van Wetenschappen, 1972, C-75, pp. 305-314.

[3] C. H. Wyndhan, W. H. Vander Walt, A. J. Van Rensburg, G. G. Rogers, and N. B. Strydom, "The Influence of Body Weight on Energy Expenditure during Walking on a Road and on a Treadmill," Int. Z. angew. Physiol. 29, 1971, pp. 285-292.

[4] R. J. Farris, H. A. Quintero, S. A. Murray, K. H. Ha, C. Hartigan, and M. Goldfarb, "A Preliminary Assessment of Legged Mobility Provided by a Lower Limb Exoskeleton for Persons With Paraplegia." IEEE Transactions on Neural Systems and Rehabilitation Engineering, MAY 2014, VOL. 22, NO. 3.

[5] R. Kobetic, M. L. Audu, G. Pinault, and R. J. Triolo, "Development of hybrid orthosis for standing, walking, and stair climbing after spinal cord injury," Journal of Rehabilitation Research & Development, Number 3, 2009, Volume 46, pp. 447–462.

[6] K. Yasuhara, K. Shimada, T. Koyama, T. Ido, K. Kikuchi, and Y. Endo, "Walking Assist Device with Stride Management System," Honda R&D Technical Review, October 2009.

 ISBN: 978-1-61208-440-4

Mobile Recruiting and Video Enhanced Job Advertisements: A Case Study Analysis within a German Job Portal

Stephan Böhm, Wolfgang Jäger, Sebastian Meurer
RheinMain University of Applied Sciences
Department of Media Management
Wiesbaden, Germany
e-mail: {stephan.boehm, wolfgang.jaeger}@hs-rm.de, s.meurer@djm.de

Ulrich Rust, Wolfgang Achilles
Jobware
Online-Service GmbH
Paderborn, Germany
e-mail: {u.rust, w.achilles}@jobware.de

***Abstract*—The proliferation of smartphones has increased significantly in recent years. Due to the availability of higher data rates, video content is increasingly being accessed via these devices. Due to changes in media usage behaviour, companies are currently faced with the challenge of adapting their recruitment activities to enable access via mobile devices and increasingly to integrate multimedia and dynamic content. In view of the above, this study examines user reactions to the enhancement of job ads with video content using a case study in a German job portal. The results of this study are promising and suggest that the integration of mobile video may be appropriate to increase interest in job advertisements. In addition, initial findings for the design and integration of such recruiting videos can be derived. The limits of panel surveys, however, also became apparent and more extensive research will be required in order to be able to make general recommendations.**

Keywords-Mobile Recruiting; Mobile Video; Job Advertisements; Video Enhanced Job Advertisements; Case Study.

I. INTRODUCTION

The adjustment of recruiting processes to changing market conditions and potential candidates' usage habits constantly presents employers and operators of job portals with new challenges. One such challenge relates to the present-day ubiquitous use of mobile devices such as smartphones and tablets. In the German market, more than 69 percent of Internet users were already accessing the Web using mobile devices in 2014. In the target group of 14 to 29 year old, this figure was even higher at over 75 percent [1][2]. The type of content consumed is also changing and the use of video content, for example, is steadily increasing. In Germany, almost every second Internet user (45 percent) aged 14 and above already accesses videos on the Internet at least once a week, while the usage rate among 14 to 29 year old is as high as 79 percent [1]. Particularly, high growth rates are evident as far as the amount of video content being accessed via mobile devices is concerned. In the US, video consumption via smartphone in 2014 increased by 73 percent compared to the previous year [3]. For the European market, a recent study from 2015 concludes that already about 60 percent of those surveyed consume short videos on their smartphones on a daily basis [4].

The need to integrate such developments –especially the shift in use to mobile devices– in corporate recruiting activities has been discussed for some time, also in international literature, under the name "Mobile Recruiting". Accordingly, mobile recruiting can be defined as "… any organizational information provided for or delivered to a mobile device in order to attract and hire potential applicants and employees."[5]. This implies continuous development and ultimately consistent mobile-optimisation in terms of technology, design and content of at least the central processing steps in the "recruiting funnel". For example, this could include the mobile optimization of career websites and job ads or even the possibility to apply directly via mobile devices. The ultimate goal of such development and optimization efforts would be to offer applicants a consistent and continuous „Mobile Candidate Experience"[6][7]. In addition to mobile-optimisation, the design of, e.g., job advertisements also needs to be adapted to the growing importance of dynamic and multimedia content. Accordingly, in a recent German study, more than 70 percent of the participating HR professionals assigned (high) importance to the integration of dynamic content such as recruiting videos in job advertisements. Besides written specifications, such as job description and profile requirements, photos and recruiting videos were identified as being important content elements of future online job advertisements [8]. A correspondingly comprehensive (mobile) recruitment strategy has, however, so far hardly been established in practice. Moreover, there are still no systematic studies that deal with the aspect of the integration of video content in (mobile-optimised) online job ads.

In view of the above, this study investigates user response to video-enhanced job advertisements using the example of a German job portal and the method of case study analysis. Currently available findings on the subject matter and relevant research are presented in Section II. The research methodology is described in Section III and subsequently, in Section IV, initial findings based on the study are presented. The paper ends with a discussion of fundamental conclusions for practical application in Section V. Furthermore, this closing section contains a discussion of the limitations of this study as well as the need for further research.

II. RESEARCH BACKGROUND

Initially, as the basis of our study and to define the research needs, the status quo with regard to mobile recruiting and the integration of video content in job advertisements are investigated using the example of the German job market. Following this, there is a brief presentation of relevant re-

 ISBN: 978-1-61208-440-4

search on mobile video to define the research needs and to set the boundaries of the research question.

A. Status Quo

According to a global recruiting study [9] in 2014 only 30 percent (34 percent) of the recruiting decision makers stated that their company provides mobile-optimized job postings (career sites). This study focuses on the German market. While mobile recruiting in the form of mobile job search is already common practice, companies often still lag behind the applicants' requirements in terms of what they offer. This is especially true for the consistent optimisation of the entire mobile recruiting process [10]. According to a recent study on the German recruiting market [11], nearly 40 percent of job seekers in 2014 already used their smartphones to search for job vacancies. About 8 percent also already apply via their mobile device. However, the optimisation of job-related content for mobile devices is still in need of improvement. According to the aforementioned study, about 70 percent of respondents report display issues on career websites when accessing them on their mobile devices. With online job advertisements on company websites, this figure rises to 75 percent. In principle, companies have, however, recognised the shift in usage habits when searching for job-related information. In a recent survey of German companies from the year 2014, 78 percent of study participants stated that the increasing use of mobile devices will have a major impact on recruiting. Nevertheless, only about 44 percent of the companies in this study reported that they provide a mobile career website. Similar to the aforementioned results of the global study, only about 31 percent of the companies in Germany mobile-optimised their job advertisements [12]. Another study, in which the recruiting offerings on the Internet of the 180 largest and most important employers in Germany [13] were examined, came to similar results. As shown in Figure 1, at the end of 2014, thus not even half of the companies included in the study provided a mobile career website, only 30 percent offered potential candidates mobile job advertisements and just one in ten companies offered the option to apply via a mobile device.

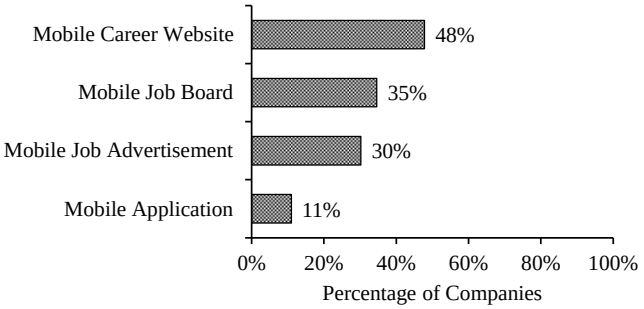


Figure 1. Mobile Recruiting Offerings in Germany [13]

Overall, job seekers in Germany still have to deal with recruiting offers which leave considerable room for improvement with regard to its mobile optimization. Therefore, many companies still need to improve the mobile experience by implementing new technologies, e.g., responsive design, to adapt their recruiting offerings to touch screens, small screen sizes and other characteristics and limitations of mobile devices. However, as mentioned above, companies are already encountering new challenges when adapting to the increased use of dynamic contents such as recruiting videos in job advertisements. As Figure 2 shows, only 10 percent of the companies in the study integrate job-related videos in their online job ads. Furthermore, mobile-optimised video job advertisements are used by just eight percent of these companies [13].

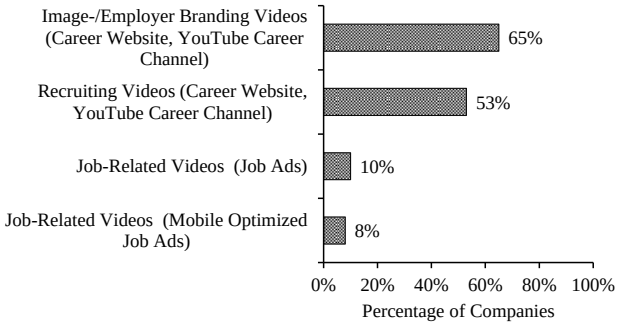


Figure 2. Implementation Status of Recruiting Videos in Germany [13]

Usage statistics of relevant YouTube channels confirm the increase in user demand for access to recruiting videos via mobile devices. According to the German job portal Jobstairs, in mid 2015 an average of 35 percent of recruiting videos on the Youtube channel were already being accessed via mobile devices. In a pilot study with mobile-optimised video-enhanced job ads, mobile access rates were as high as 40 percent [14]. Against this background, this paper focuses on video-enhanced job advertisements as a next level of mobile recruiting.

B. Related Work

A number of research studies can already be found on the structure and content of job advertisements. Relevant articles, however, often concentrate on the content analysis of these ads to identify requirements and characteristics in terms of job descriptions or professional fields [15][16][17]. A closer focus on requirements, usability aspects and design principles specifically of mobile job advertisements can be found in [7][18]. In these papers, however, specific aspects of the integration of video content in job advertisements have not yet been considered. The issue of mobile video has, however, already been investigated more comprehensively in papers dealing with the topic separately from the context of job advertisements and recruiting. These include various analyses dealing with the aspects of technical implementation of mobile video services or the perception of their quality and their market perspectives [19][20][21]. Research can also be found which deals with the use of (mobile) video for marketing and advertising purposes [22][23]. This research offers few findings with regards to the integration of recruiting videos in job advertisements. To our knowledge, the importance and the benefits of video-enhanced job ads have thus not yet been examined in any research papers. More specific recommendations on the content and user-oriented

 ISBN: 978-1-61208-440-4

design of video-enhanced job ads in the context of stationary and mobile use are also nowhere to be found. In view of the above, this study attempts to present initial findings to close this research gap. The focal point of the investigation in this case is the examination of user reactions that arise when presented with video-enhanced jobs ads.

III. APPROACH

The placement of video content in mobile job ads represents a very specific application of mobile video. In order to investigate the potential of enhancing mobile job ads with recruiting videos appropriately, the analysis must be applied to suitable applicant target groups. Not only do these target groups need to match the requirements for the advertised job, they must also be interested in applying for the position. In light of these facts, it can be assumed that limited findings regarding interest in and reception of video content in job ads, can be obtained by using a standard online survey and a fictitious job ad.

For this reason, this study is based on an analysis of real job ads that were posted on the German job portal *Jobware*. A systematic variation of characteristics of the job ads was therefore not possible. However, the case studies provide first insights into the interest in and the reception of video-enhanced (mobile) job ads. As a first step, three companies were recruited as partners that agreed to integrate recruiting videos in their job ads and to permit the analysis of usage data. These were businesses ranging from a medium-sized "hidden champion" to a major global player. Basic characteristics of these companies, which cannot be named directly due to the confidential nature of the study, and of the videos used, are shown in Table I.

TABLE I. CASE STUDY RECRUITING VIDEO CHARACTERISTICS

	Video 1	Video 2	Video 3
Industry	Electronics	Insurance	Automotive
Enterprise Size (Employees)	~ 14.000	~ 147.000	~ 280.000
Enterprise Reputation	Hidden Champion	Blue Chip	Global Player
Position	Engineer	IT-Trainee	Engineer
Video-type	Job-related	Job-related	Image Video
Length	1:59 min	1:46 min	7:02 min

The differences in the type and length of the video are due to the fact that only two of the companies provided or produced a job-related video for the study. In the case of the third company, an existing recruiting/image video was used. Due to the cooperation with the job portal, it was possible to carry out four different sub-analyses for each of the job ads described:

- *Job Ad Analytics*: The statistics on retrieval of the job ad in the job portal form the basis of the subsequent analysis steps.
- *Video Analytics*: In addition to the number of video views, the system could also detect how often each video was watched up to 25%, 50%, 75% or up to the end.
- *Online Questionnaire*: After exiting the video in the job ads, users of the job portal were asked to voluntarily fill in a short questionnaire.
- *Online Panel*: A survey was carried out in an online panel outside the job portal.

In the online panel, participants were asked to select one of the three jobs in accordance with their areas of interest and to fill in a questionnaire after viewing the accompanying videos. The results of the panel survey were intended to supplement and verify the previously obtained responses given by job seekers in the job portal.

IV. RESULTS

After recruiting the cooperation partners, the video-enhanced (mobile-optimised) job ads were posted on the job portal. The data for the job ads on the job portal was acquired for the period May 25th to June 22nd, 2014. The panel data were collected between June 20th and 27th, 2014. Table II shows the views of the job ads included in the study. The number of visitors to each job advertisement shows that the number of views significantly varies according to the size of the associated company. Presumably this is also an expression of either the company's prominence or the attractiveness of the position. Surprisingly, however, the click-through-rate from the job ads to the companies' career websites and application forms varied considerably. The click-through rate from the job ad to the career website of the lesser known company ("hidden champion") was almost three times as high as that of the global player. Furthermore, the click-through rate to the application form was more than twice as high for the hidden champion. Target groups and advertised vacancies in job ad 1 and job ad 3 were similar. The differences in click rates could therefore be seen as a first indication that job ad 1 was more able to arouse interest in not only the company but also applying for the job.

TABLE II. JOB ADVERTISEMENT ANALYTICS

	Job Ad 1	Job Ad 2	Job Ad 3
Job Ad Visitors	182	263	287
Click-through Rate (Clicks) to Career Website	15% (28)	8% (21)	5% (13)
Click-through Rate (Clicks) to Application Form	19% (35)	10% (25)	8% (22)

When analysing the video views shown in Table III, a similar picture emerges. The hidden champion's video 1 had by far the highest play rate and was played during almost every tenth view by the users of the job portal. The difference is even more significant when considering the extent to which the video was actually played. Although approximately 24 percent of viewers discontinued video 1 a quarter of the way through, this discontinuation rate is nearly 63 percent for video 3. Only about 17 percent watched video 3 to the

 ISBN: 978-1-61208-440-4

end, whereas 56 percent of the viewers did so with video 1 and 42 percent with video 2. Here, the length of the video is likely to have had a significant effect. Both videos 1 and 2 were under 2 minutes while video 3 ran over 7 minutes. It is also interesting that although the two videos of companies 1 and 2 were similar in length and both job-related, they were played at very different rates of frequency. This could indicate that videos with job-related content and those with company-related content arouse different levels of interest or offer varying levels of benefit. As already mentioned, company 1 is not a very well-known company, while company 2, as a blue chip, is likely to be known by almost all the applicants. Thus, this may mean that applicants could have a higher need for information in the case of company 1.

TABLE III. RECRUITING VIDEO ANALYTICS

	Video 1	Video 2	Video 3
Play Rate (overall)	9.5%	2.5%	2.2%
Completion Rate (minimum played video length)			
25%	23.9%	35.8%	62,7%
50%	8.5%	13.4%	16.9%
75%	11.3%	9.0%	3.4%
100%	56.3%	41.8%	16.9%

As mentioned earlier, an online questionnaire was implemented within the job portal, which could be answered voluntarily on exiting the job ad. However, the job seekers' willingness to participate was very low (job ad 1: 15; 2: 22; 3: 12). In addition, the participants did not always answer all the questions in the questionnaire. In total, the questionnaire included five questions. The participants were first asked whether they had even watched the video in the job posting (Job ad 1: 87%; 2: 68%; 3: 75%). Those that viewed the video were then asked about the authenticity and quality of the video. In addition, they were asked whether the video presented a better picture of the job and whether they generally considered videos in job ads to be useful. In this case, the respondents could answer in the form of a 5-point Likert scale (degrees of agreement). The results of this online survey are shown in Figure 3.

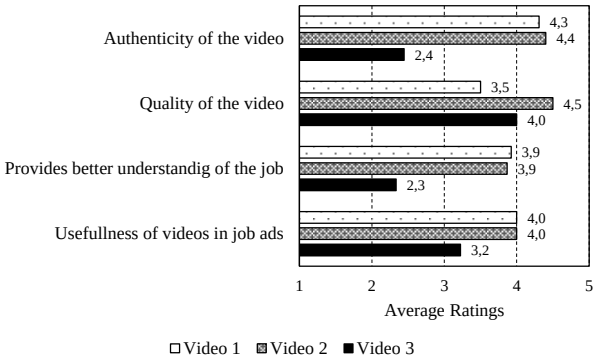


Figure 3. Jobportal Results on the Online Questionnaire

Of particular note is the fact that the results are very similar for the job-related videos 1 and 2. Only the quality of video 2 is rated as being significantly higher than that of video 1. The ratings for video 3, which, in contrast to the other two, was a recruiting video more focused on image and employer brand, are however significantly different. Despite high quality ratings, the authenticity was rated as very low and the job seekers were not able to obtain a better picture of the advertised position. We can therefore assume that the added value of this video in the job ad was perceived as rather low, which may have also had an effect on the otherwise high rating of the fundamental usefulness of videos in job advertisements.

A final survey within an online panel with a total of 100 participants was intended to obtain additional information on the reception of the recruiting videos in the job ads and also to verify the results of the online questionnaire within the job portal. As a first step, the survey participants were asked to choose a suitable job advertisement and subsequently asked whether the video had also been viewed. The play rate identified here (video 1: 31%; 2: 32%; 3: 19%) turned out significantly higher than in the video analytics of the job portal, which is probably due to the test situation. Nevertheless, we can see here again that video 3 was accessed much less often. Subsequently, those who had not watched the video were asked for their reasons. In Figure 4, we can see that the importance of the reasons is similar across all the videos. The most common reason given was that the video had not been seen or found. Therefore, references to the video as well as its positioning in the ad must be improved. In this regard, only the visibility of the video in job ad 2 is better.

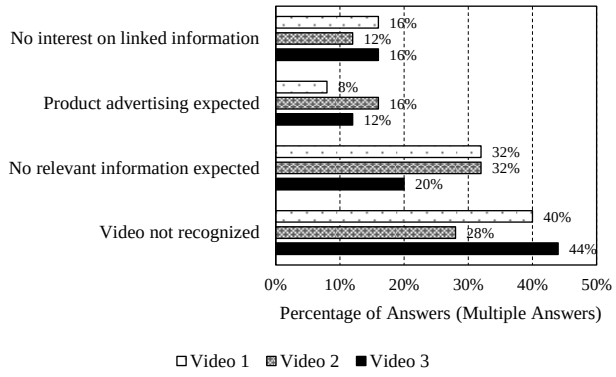


Figure 4. Panel Results on the Reasons for Not Watching the Video

With regards to the reasons for discontinuing viewing, the study participants' presumptions of product advertising in the hidden champion's job ad and of non-relevant content in the global player's job ad turned out to be below average. This is likely to be connected to the differences in terms of (brand) awareness.

Finally, three of the questions already contained in the online survey on the job portal were posed to the participants of the panel survey. It should be noted that all participants in the panel survey (job ad 1: 35; 2: 34; 3: 31) were asked to watch the video in the job ad in order to answer these questions. The results of this survey are shown in Figure 5. The

 ISBN: 978-1-61208-440-4

question on the subject of quality was no longer asked as no additional findings were expected in the test situation of the panel. Overall, the results of the panel survey were rather disappointing and characterised by a significantly lower range of variation and less clear statements than the results of the online survey on the job portal. On the one hand, this could be due to the fact that the panel participants were randomly selected for the survey. Although screening questions were asked on the field of study and other socio-demographic characteristics, the respondents were not directly in the process of looking for a job.

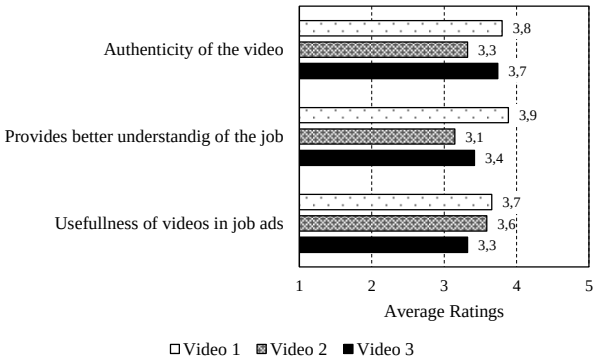


Figure 5. Panel Results on the Questions of the Online Questionnaire

Additionally, in the second part of the survey, participants were prompted to look at the job ad regardless of interest. Thus, brand awareness or the professional appearance of the video may have played a role in the evaluation, rather than applicant-specific information needs.

V. CONCLUSIONS

The current shift in Internet usage from stationary to mobile devices is a major challenge to companies to adapt their offerings in the area of e-recruiting. Online job ads, career websites and application management systems have to be optimised for access via mobile devices. While this initially affected primarily the presentation and layout of the user interface design, employers must now also increasingly adjust the content design of their offers. The integration of recruiting videos in job ads, for example, is a possibility here. Especially, when considering that many job seekers are now accessing the information via mobile devices, video will allow the presentation of more complex content and context in a mobile friendly manner. In this study, the integration of such videos in job advertisements was systematically investigated. As a result, in particular the following findings can be noted:

- *Usage*: Applicants generally consider additional video content in job ads useful. In the study, however, click rates of under 10 percent were identified.
- *Video duration*: The length of the videos should be less than two minutes. Longer videos lead to significantly higher discontinuation rates.
- *Video content*: Job-related videos that convey an authentic picture of the advertised vacancy are preferable to general corporate image videos and content related to employer branding.
- *Embedding*: The videos should be directly referred to in the job advertisement, also in order to clearly differentiate them from other advertising. Videos must be positioned directly in the job advertisement to make them easy to find.

Based on the results of this study, we can also assume that especially lesser known companies could benefit from video integration. While applicants probably already have some knowledge of companies with a well-known employer brand, lesser known employers can convey this information via an appropriate video.

The generalisability of the findings of the present study is limited. Firstly, the fact that the investigation was carried out as a case study is significant, as only three different job ads could be investigated and interventions in the ads were limited to the integration of the video. Furthermore, the system was not yet able to determine the share of observations directly attributable to mobile access. Such enhancements in the analysis should accordingly be the subject of further research. This could be done, for example, on the basis of structured A/B testing. The results of the panel surveys show that an integration of the analysis in real job portals is preferable in order to evaluate the attitudes and behaviour of real job seekers. In addition, differences in the content reception of video content in stationary and mobile contexts should be examined in further studies. Important insights could also be derived from such studies as to how the design and the contents of recruiting videos can be suitably adapted to mobile usage situations.

REFERENCES

[1] ARD/ZDF, 2014, ARD/ZDF-Onlinestudie 2014. http://www.ard-zdf-onlinestudie.de/, retrieved: October, 2015.

[2] Initiative D21, 2014, Mobile Internetnutzung 2014. http://www.initiative21.de/wp-content/uploads/2014/12/Mobile-Internetnutzung-2014_WEB.pdf, retrieved: October, 2015.

[3] Brightroll, 2014, Mobile Video Advertising Strengthens TV Media Investments. http://info.brightroll.com/rs/brightrollinc/images/BrightRoll_mobile_video_study.pdf, retrieved: October, 2015.

[4] On Device Research, 2015, Mobile Video 2015: A global perspective. http://www.iab.net/media/file/IAB_Mobile_Video_Usage_FINAL.pdf, retrieved: October, 2015.

[5] S. Böhm and S. J. Niklas, "Mobile recruiting: Insights from a survey among German HR managers," SIGMIS-CPR '12 Proceedings of the 50th annual conference on Computers and People Research, Milwaukee, USA, May 31-June 02, 2012, 2012, pp. 117–122.

[6] S. Böhm, "Behavior and expectations of mobile job seekers: An industry study focusing on job boards," SIGMIS-CPR '13 Proceedings of the 2013 Annual Conference on Computers and People Research, Cincinnati, USA, May 30-June 01, 2013, 2013, pp. 105–110.

 ISBN: 978-1-61208-440-4

[7] S. Böhm, W. Jäger, S. Meurer, and S. J. Niklas, "Optimizing the Usability of Mobile Job Advertisements: A Multi-Method Approach," International Journal on Advances in Intelligent Systems 7, 1&2, 2014, pp. 257–266.

[8] Jobstairs, 2015, Die Zukunft der Online-Stellenanzeigen ist mobil und dynamisch. http://www.jobstairs.de/index.php?ac=page&ci=1004, retrieved: October, 2015.

[9] Linkedin, 2015, 4th Annual Report: 2015 Global Recruiting Trends. https://business.linkedin.com/talent-solutions/resources/recruiting-tips/global-recruiting-trends, retrieved: October, 2015.

[10] S. Meurer, W. Jäger, and S. Böhm, 2014, Mobile Recruiting 2013. Eine empirische Untersuchung zur Bewerberansprache über mobile Endgeräte. http://www.djm.de/mr2013/mobile-recruiting-2013_bericht.pdf, retrieved: October, 2015.

[11] T. Weitzel, A. Eckhardt, S. Laumer, C. Maier, A. von Stetten, C. Weinert, J. Wirth, and B. Kraft, Bewerbungspraxis 2015. https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/wiai_lehrstuehle/isdl/Bewerbungspraxis_2015.pdf, retrieved: October, 2015.

[12] T. Weitzel, A. Eckhardt, S. Laumer, C. Maier, A. von Stetten, C. Weinert, J. Wirth, and B. Kraft, Recruiting Trends 2015. https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/wiai_lehrstuehle/isdl/Recruiting_Trends_2015.pdf, retrieved: October, 2015.

[13] W. Jäger and S. Meurer, Human Resources im Internet 2014. RheinMain University of Applied Sciences (forthcoming), 2015.

[14] Jobstairs, Expert interview on (mobile) recruiting video usage on Youtube carrer channel, 19.06.2015.

[15] B. K. Addom, Y. Kim, and J. M. Stanton, "eScience professional positions in the job market," Proceedings of the 2011 iConference, February 8-11, 2011, Seattle, USA, pp. 630–631.

[16] C. Lauer and E. Brumberger, "Architect, Developer, Designer," Proceedings of the 32nd ACM International Conference on The Design of Communication, September 27-28, 2014, Colorado Springs, USA, pp. 1–2.

[17] F. Steinmann, K.-I. Voigt, and T. Schaeffler, "Engineering Competences: A Content Analysis of Job Advertisements," Proceedings of PICMET '13: Technology Management for Emerging Technologies. July 28-Aug 1, 2013, San Jose, USA, pp. 1925–1929.

[18] S. Böhm, S. Niklas, and W. Jäger, 2015, "A Multi-Method Approach to Assessing the Usability of Mobile Job Advertisements (Working Paper)," Proceedings of CENTRIC 2013, The Sixth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services, October 27-November 1, 2013, Venice, Italy, pp. 108–111.

[19] K. J. Ma, R. Bartos, S. Bhatia, and R. Nair, "Mobile video delivery with HTTP," IEEE Commun. Mag. 49, 4, 2011, pp. 166–175.

[20] J. Xue and C. W. Chen, "A study on perception of mobile video with surrounding contextual influences," 2012 Fourth International Workshop on Quality of Multimedia Experience (QoMEX 2012), pp. 248–253.

[21] A. Dillistone, R. Moreno, and C. Voisey, "Market perspective," MoViD '10 Proceedings of the 3rd Workshop on Mobile Video Delivery, October 25, 2010, Firenze, Italy, p. 69.

[22] S. S. Krishnan and R. K. Sitaraman, "Understanding the effectiveness of video ads," Proceedings of the 2013 Conference on Internet Measurement, October 23-25, 2013, Barcelona, Spain, pp. 149–162.

[23] M. Carpen Amarie, I. Pefkianakis, and H. Lundgren, "Mobile video ad caching on smartphones," UbiComp '14 Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, September 13 - 17, 2014, Seattle, USA, pp. 57–61.

 ISBN: 978-1-61208-440-4

User-oriented Product Information Management with Semantic Technologies

Bastian Eine, Werner Quint
Faculty of Design – Computer Science – Media
RheinMain University of Applied Sciences
Wiesbaden, Germany
bastian.eine@hs-rm.de, werner.quint@hs-rm.de

***Abstract*— Product information is important for companies' employees and customers and thus, for the companies' overall success. There is the need to manage and represent complex product information in a way that addresses different user capabilities and perspectives because individuals may have different information needs and different perspectives on product information. This paper describes the idea that product information can be represented and managed user-oriented with the help of semantic technologies.**

Keywords-product information management; user-orientation; semantic technologies.

I. INTRODUCTION

Today, business transactions, including business processes and acquisition of data, are almost entirely electronic in nature. With regards to product information, enterprises are faced with the challenge of how to handle more and more information about products [1]. Adding to the challenge is customer demand for mass customization. This requires the management of even more information about relations between product parts and product features. Information systems may help to reduce the complexity of product information for users, but for this the information systems need functions based on specific technologies. Accordingly, this paper describes the idea of representing and managing product information user-oriented with the help of semantic technologies. In Section 2, product information management (PIM) as a part of enterprise information technology will be described. Section 3 shows challenges for companies that want to apply PIM. The idea of using semantic technologies for PIM will be discussed in Section 4. Finally, Section 5 describes the approach of how to develop a solution.

II. PRODUCT INFORMATION MANAGEMENT

Internally, information about products is not always consistent. Product information is created by specialized departments, or its employees, which have different perspectives on information [2]. Also, product information is created outside of a company, managed, passed on and taken over both by other companies, e.g., suppliers, manufacturers and by individuals like customers [3]. If a company stores and manages the different information about products in a central and media neutral way, the expense of maintaining, searching for and presenting product information can be reduced significantly [4]. Thus, the company will also able to reuse product information consistently in different departments. Furthermore, consistent product information can reduce the number of customer requests and wrong orders and deliveries.

The establishment of electronic business processes brought a number of improvements for companies, e.g., automatic handling of purchasing and selling products. Therefore, data and information about products are almost exclusively managed in companies' information systems [5]. That being said, not all information systems have the ability to fully encompass the PIM needs of a company. For example, Enterprise Resource Planning (ERP) systems are set up to represent all business processes of a company in order to increase the overall cost-effectiveness [6]. Often, ERP systems do not include product marketing descriptions, product pictures or complementary technical product data. Additionally, there are a number of other types of information systems, with which product information can be managed, e.g., Content Management Systems or Customer Relationship Management Systems. Finally, PIM systems provide centralized and media neutral data storage, data management, and data output of product information [7]. PIM systems offer a number of functions, which enable companies to use product information consistently inside and outside of the company. For instance, PIM systems can assist companies and employees with product classification, translation management, media asset management and data output to different media (e.g., print catalogue, technical data sheets, online shop, or other third-party systems) [4].

III. CHALLENGES

One challenge for managing product information is ensuring a consistent description and interpretation of product information by users [8]. Therefore, it is necessary to implement syntactical, as well as semantical restrictions or defaults to avoid redundant statements or wrong interpretations. Another challenge is to capture and represent complex product information, e.g., relations between products or product features. When composing a configurable product it has to be taken into account that selecting a certain product feature may exclude other product features. This information is essential for

 ISBN: 978-1-61208-440-4

employees, who are responsible for maintaining product information but also for customers, who want to configure a product or satisfy their information need. Accordingly, it is necessary to capture, manage and present the information in a clear and easy to understand manner. PIM systems are often advertised as being able to manage a large variety of relations between products and product features. Also, they should be able to present the products and relations between product features in different media easily. This shows that there might be alleged solutions already with which product information can be gathered and displayed in individual formats. However, it is unclear, on which technology and approaches the providers are basing their PIM systems. Also, it has not yet been analyzed whether existing solutions take different user perspectives into account and if they do, how satisfying are they.

IV. SEMANTIC TECHNOLOGIES

The use of semantic technologies, when creating and representing information and relations between concepts, can help interpret complex information by identifying the corresponding context. Semantic technologies can make it easier to understand the meaning and purpose of complex concepts as well as share knowledge for humans and machines [9]. For PIM, semantic technologies can be based on simple approaches like glossaries (lists of words and their definitions), taxonomies (hierarchies for terms) and thesauri (relations of similarity and synonyms) to avoid syntactical and semantic problems when creating and interpreting product information. Approaches with more semantic richness are topic maps [10][11] and ontologies [12]. Besides describing product information these two semantic technologies can be utilized to also capture and represent the relations and connections to other products, product components, product features and further product information. With ontologies it is also possible to represent rules which are associated with the relations mentioned before. Hepp [13] and Brunner et al. [14] are examples for describing and linking products and services with the help of ontologies.

Normally, the complex labels and structures of semantic technologies are created with complex tools (high range of functions, programming skills necessary). It cannot be expected that users, who are responsible for maintaining product information in a company, are able to easily understand and utilize these structures and tools. Even more problems can arise when a user has to modify or customize the structures of the implemented semantic technologies. For example, when a product taxonomy has to be modified, because product lines are merged. Thus, a user not only has to get to know the tool for modifying but also has to understand structures and relations of a product taxonomy. Only then will the user be able to make useful modifications without causing errors, inconsistencies or contradictions. This shows that employees and customers can handle complex product information more easily if structures and relations of semantic technologies for PIM are represented in an understandable and user-oriented way.

V. APPROACH

The idea of this paper is to utilize the concept of semantic technologies for user-oriented representation and management of product information. First step will be a literature review to provide an overview of the current concepts in research and the research needs in the context of user-oriented representation of complex product information and semantic technologies. Also, companies using PIM will be interviewed for user requirements analysis. Next steps will be to develop a method and a software prototype to support user-oriented PIM. This method and prototype will be evaluated with the help of companies using PIM.

REFERENCES

[1] McKinsey Global Institute. Big data: The next frontier for innovation, competition, and productivity. [Online]. Available from: http://bigdatawg.nist.gov/MGI_big_data_full_report.pdf 2015.06.01.

[2] J. Stark, Product Lifecycle Management - 21st Century Paradigm for Product Realisation. London, Springer, 2011.

[3] B. Otto and M. Ofner, "Towards a Process Reference Model for Information Supply Chain Management," The 18th European Conference on Information Systems (ECIS 2010), Jan. 2010, pp. 1-12.

[4] T. Lucas-Nülle, Product Information Management in Deutschland - Marktstudie 2005. Mammendorf, Pro-Literatur-Verl. Mayer-Scholz, 2005.

[5] J. A. O'Brien and G. Marakas, Management Information Systems. New York, McGraw-Hil/Irwin, 2011.

[6] M. Sumner, Enterprise Resource Planning. Indiana University, Prentince Hall, 2005.

[7] Gartner, Magic Quadrant for Product Information Management 2007. [Online]. Available from ftp://public.dhe.ibm.com/software/emea/de/db2/Gartner_MDM_Magic_Quadrant_PIM.pdf 2015.06.01.

[8] R. Y. Wang and D. M. Strong, "Beyond Accuracy: What Data Quality Means to Data Consumers" J. Manag. Inf. Sys., vol. 12, Mar. 1996, pp. 5-33.

[9] T. Pellegrini and A. Blumauer, "Semantic Web und semantische Technologien" in Semantic Web - Wege zur vernetzten Wissensgesellschaft, T. Pellegrini and A. Blumauer, Eds. Berlin: Springer-Verlag, pp. 9-25, 2006.

[10] S. Pepper, The TAO of Topic Map - Finding the Way in the Age of Infoglut. [Online]. Available from: http://www.ontopia.net/topicmaps/materials/tao.html 2015.05.13.

[11] L. M. Garshol, "Metadata? Thesauri? Taxonomies? Topic Maps! Making Sense of it all". J. Inf. Sci., vol. 30, August 2004, pp. 378-391.

[12] T. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Sharing" Int. J. Hum.-Comp. Stud., vol. 43, Nov. 1995, pp. 907-928.

[13] M. Hepp, "GoodRelations: An Ontology for Describing Products and Services Offers on the Web" The 16th Int. Conf. Knowl. Engin. Knowl. Manag. (EKAW2008), Springer LNCS, Sept. 2008, pp. 332-347.

[14] J.-S. Brunner et al., "Explorations in the use of semantic web technologies for product information management" 16th Proc. Int. conf. www (WWW2007), ACM, May 2007, pp. 747-756.

 ISBN: 978-1-61208-440-4