CENTRIC 2011

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

ISBN: 978-1-61208-167-0

October 23-29, 2011

Barcelona, Spain

CENTRIC 2011 Editors

Karla Felix Navarro, University of Technology Sydney, Australia

CENTRIC 2011

Forward

The Fourth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services (CENTRIC 2011), held on October 23-29, 2011 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 2011 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 2011. 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 2011 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 2011 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 hope Barcelona provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

CENTRIC 2011 Chairs

Advisory Chairs

Vitaly Klyuev, University of Aizu, Japan

Mario Hoffmann, Fraunhofer-Institute for Secure Information Technology (SIT) - Darmstadt, Germany

Guadalupe Ortiz, University of Cádiz, Spain

Research Institute Liaison Chairs

R. Hariprakash, Bharath Institute of Higher Education &. Research - Chennai, India

Muhammad Iftikhar, University of Malaysia Sabah, Malaysia

Kong Cheng, Telcordia Technologies - Piscataway, USA
Gottfried Zimmermann, Eberhard Karls Universität Tübingen, Germany

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

CENTRIC 2011

Committee

CENTRIC Advisory Chairs

Vitaly Klyuev, University of Aizu, Japan
Mario Hoffmann, Fraunhofer-Institute for Secure Information Technology (SIT) - Darmstadt, Germany
Guadalupe Ortiz, University of Cádiz, Spain

CENTRIC 2011 Research Institute Liaison Chairs

R. Hariprakash, Bharath Institute of Higher Education &. Research - Chennai, India
Muhammad Iftikhar, University of Malaysia Sabah, Malaysia
Kong Cheng, Telcordia Technologies - Piscataway, USA
Gottfried Zimmermann, Eberhard Karls Universität Tübingen, Germany

CENTRIC 2011 Industry/Research Chairs

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

CENTRIC 2011 Technical Program Committee

Hasan Ibne Akram, Technische Universität München, Germany
Sarfraz Alam, University Graduate Center (UNIK) - Kjeller, Norway
Giner Alor Hernandez, Instituto Tecnologico de Orizaba, México
Atta Badii, University of Reading, UK
Rémi Bastide, Université Champollion, France
Carlos Becker Westphall, Federal University of Santa Catarina, Brazil
Nadia Bennani, INSA-Lyon - Villeurbanne, France
Mahmoud Boufaida, Mentouri University of Constantine, Algeria
Mahmoud Brahimi, University of Msila, Algeria
Gunnar Brink, Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Germany
Marco Bruti, Telecom Italia Sparkle S.p.A. - Roma, Italy
Sylvie Calabretto, INSA de Lyon - Villeurbanne, France
Kong Cheng, Telcordia Research, USA
Max Chevalier, Université de Toulouse, France
Jeffrey J. Evans, Purdue University, USA
Rainer Falk, Siemens AG - München, Deutschland
Stenio Fernandes, Federal University of Pernambuco, Brazil
Agata Filipowska, Poznan University of Economics, Poland
Ulrich Flegel, University of Dortmund, Germany
David Gomez-Barquero, Universidad Politecnica de Valencia, Spain
Vic Grout, Glyndwr University, UK

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

Methodolgy for Bridging the Brain and Body for Briliant Innovations

Maher Sefein., Muhammad Iftikhar & Osman Ali
School of Medicine
Universiti Malaysia Sabah
Kota Kinabalu, Malaysia
e-mail: {maher, ifti, osmanali}@ums.edu.my

Muralindran Mariappan
School of Engineering & Information Technology
Universiti Malaysia Sabah
Kota Kinabalu, Malaysia
e-mail: murali@ums.edu.my

***Abstract*—Medical doctors are likely to be clinically oriented due to monitory benefits and busy life schedule; therefore this reduces their interest in research and innovation. In order to increase the innovation interest, the School of Medicine, Universiti Malaysia Sabah has initiated a new strategy to fasten the research and innovation by CLIP Method. This concept is based on establishment of Creative Work Station, learning enhancement by Mechatronics Workshop, collection of innovative ideas and designing new projects. Preliminary results are given and further evaluation is required to establish the concept. Initial result shows that experienced physician's brain (knowledge & ideas) can be utilized by hands-on techniques (body activities) to get more creative ideas leading to patentable projects.**

Keywords-creativity; research; innovation; medical; prototype

I. BACKGROUND

Creativity is a key component to the long-term survival of a company [1]. Though creativity is popularly considered as in-born qualities of a genius, they can be nurtured and developed through continuous training [2]. An undergraduate senior level courses which were designed by University of Florida, "Creative Engineering I & II Sequence" were able to help in their institutions by training [3][4][5]. Despite the fact that there has been significant improvement in the healthcare industry, inefficiency still exists and little is accomplished in understanding how to overcome those inefficiencies using innovation in healthcare [6].

Creativity is a problem identification and idea generation whilst innovation is an idea selection, development and commercialization [8]. Innovation can therefore be seen as the process that renews or improves something that exists and not, as is commonly assumed, the introduction of something better [9]. All innovation begins with creative ideas and successful implementation of creative ideas within an organization leads to innovations [10]. In this view, creativity by individuals and teams is a starting point for innovation; the first is necessary but there is no sufficient condition available for the second [11]. Tidd *et al* [12], stated that there are four types of innovation; consequently the innovator has four pathways to investigate when in search for good ideas i.e. product, process, positioning and paradigm innovation. Innovation in healthcare is defined as those changes that help healthcare practitioners to focus on the patient by helping healthcare professionals to work smarter, faster, better and with more cost effectiveness [1].

Based on survey of 775 healthcare professionals from the US, UK, Germany and India [13]. In this it was stated that the problem is not a lack of ideas [1]. Difficulties lie in the diverse blockages to new ideas finding their way into widespread and transformative change [13]. On-the-job-training to keep healthcare executives and employees up to-date would increase their familiarity with technology and also enhances their likelihood of adopting new information technology [13] [37] [38]. This study summarizes all these facts as given below.

- Lack of ideas is not a problem.
- On job-training will enhance their likelihood of adopting new IT and generation of brilliant ideas.
- Lack of sharing ideas is always present therefore capacity to generate ideas can be enhanced by sharing.
- Sharing attitude will make project selection easier which leads to potent commercialise product.
- Another road block for doctors' innovation is lack of the venue whereby they can work on their innovation in hospitals or medical institutes.

CLIP Methodology is developed to overcome these road blocks. It is a strategy that bridges the brain (which is full of ideas) and the body (hand' skills) to come out with brilliant innovations. CLIP is designed from first word of the four components: **C**reative Work Station, **L**earning Skills, **I**deas Generation and **P**rototype Production. Therefore in CLIP Methodology, (C) is the venue, where learned skills (L) work together , to transform (I) into physical Protypes (P) .

II. OBJECTIVES

In view of the availability of different expertise in Universiti Malaysia Sabah (UMS) and the available facilities in School of Medicine, it was found that CLIP concept can circumvent the above mentioned problems that hinder medical practitioner's innovations. The comparison of creativity, innovation and CLIP methodology is summarized in "Fig. 1". The methodology focuses on idea creation, generation and selection. A provision of basic structure is required for the creation of ideas among medical personnel who were never involved in innovative activities. Therefore innovation directly starts from idea selection and creativity which is simply problem identification and then followed by idea generation. CLIP methodology creates comprehensive approach to produce prototypes projects leading to patentable products.

 ISBN: 978-1-61208-167-0

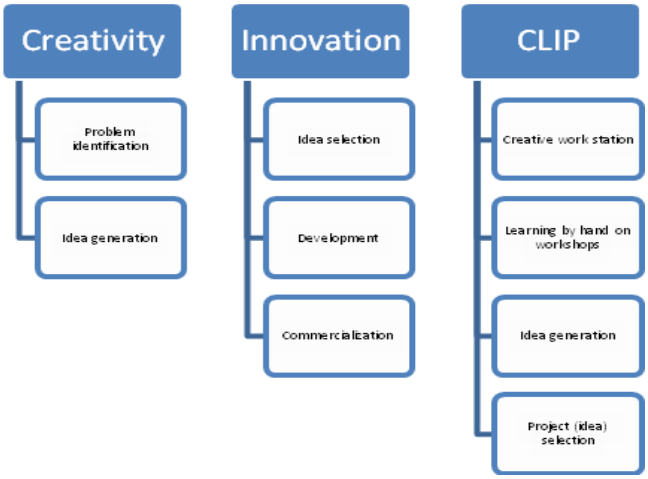


Figure 1. Comparison of Creativity, Innovation & CLIP

All school of medicine members are encouraged to join CLIP activities. They are referred here as MAST , where (M) stands for Management staff,(A) for Academicians , (S) for Students and (T) for Technical staff .

CLIP enhances the creative thinking and innovative attitude among MAST through various non-medical training programms [29][37]. The training programs (L) in CLIP apply the theory of multiple intelligences by Dr. Gardner [29]. Accordingly different potential pathways are chosen in (L) sessions, whereby the first Mechatronic workshop is to enhance the {Spatial intelligence} [30]. In other work, a pharmacy school taught electrocardiogram basics by using dance and movement {Musical intelligence}. In that trial the experiment was effective and had good learning outcome [31]. The technological arm along with the medical staff will narrow the gap between the practicing doctors and the technology which will definitely improve the practice and also facilitate innovating process. This will build a background of physical understanding and exemplary human physiology, which is beneficial for both the students and lecturers. Meanwhile, the engineers and other expertise who collaborate in such training (L) will get a new sense of understanding of the human body mechanisms [31]. The reciprocal discussion and sharing of experience will enhance coloration, which will impact on the development of new applications in robotics and biomedical engineering. CLIP will also enhance inter-university cooperation with other schools and institutes in UMS e.g., School of Engineering & IT, Biotechnology Research Institute and School of Science & Technology, Fine arts, Social science, etc.

Valuable trainings (L) and good facilities (C) will encourage MAST to participate in innovation competitions at the university and national levels. When the students (S) participate in the CLIP, they will have a better understanding of the rule of technology in the medicine and to improve learning process [14] . They also can use the technology to execute learning projects as anatomy models [35]. It is an important factor in their learning process, as 90% is retained in the memory if it is practically done. Although it is controversial, it is still commonly refereed percentage as the Learning Pyramid which was first published by an employee of Mobil Oil Company in 1967 in a Film and Audio-Visual Communications [13] [14][35].

Having the facilities to build up own protypes (P), will cut the cost, time and overcome the lack of facilities which hampers many of innovative ideas (I). These processes will enable the School of Medicine to have many prototypes which will support one of its main objectives.

III. MATERIALS AND METHODS

Wherever the targets are, a clear methodology is dependent on the available resources. Creative work station, Learning, Innovative ideas and Projects (CLIP Method) are its four components as described below:

A. *Creative work station (C)*

The need for it was recognized in the year 2008 when the authors started to transform the innovative ideas and projects into prototypes. It was initially difficult to develop it in the School of Medicine, as there was no venue for such type of hardware activities. During this period, their own facilities was used to for development. The same problem was also faced by many of MAST members, especially those who had no personal equipments to fabricate the prototypes. So having a work station would enable the user to transform the birillent ideas into feasible innovations. Creative Workstation started on May 2010. The committee which look after the (C) is the school's dean , and membership of 5 talented members from MAST . An academic member from School of Engineering and IT joined the committee as a technical consultant.
A spacious room was allocated as (C). The venue was renovated to support the developing basic hardware work.

B. *(L) activity of CLIP:*

The first mechatronic workshop was an introductory to the foundation for the mechanics and electronics.

The training was led by a lecturer from School of Engineering and supervised by 3 postgraduate students as technical assistants. The participants were divided into 3 smaller homogeneous groups. Among these groups, each pair had to accomplish their own mini project.

Robotic Research Laboratory in School of Engineering was selected as the venue. The wrkshop was held for two successive days.

Training Structure: The technical skills (L) are ranked in different levels from the basic to advance. All participants should first join the basic mechatronic workshop to be exposed to basic technological skills.

 ISBN: 978-1-61208-167-0

The planned knowledge and skills were managed to be integrated with the attitude (Schön, 1983) of the participants to achieve the goal [37].
The workshop had 4 main themes as given below.

a) *Safety of mechanical tools and electronic devices*
b) *Basic electronics and circuits.*
c) *Basic sensors.*
d) *Basic motor and control.*

The objective of this introductory workshop was to expose the participants to the basic skills in handling different materials. They were exposed to wood and metal work. The basic wood work skills were (drilling, cutting with different saws, gluing and measuring). While in metal work, they practiced (cutting with different saws, metal drilling, riveting). This also later covered the basic electronics skills (soldering, reading circuits, identifying components, testing current and voltage) [39] [40]. Each theme is covered through a one hour lecture followed by 3 hours of Hands-On session. The four lectures enabled the participants to understand the basics, application of the electronics and mechanical integration gradually in a comprehensive way. The outline of the workshop is shown in Table 1.

TABLE 1. THE OUTLINE OF 1ST BASIC MECHATRONIC WORKSHOP TRAINING

Theoretical	Hands-On		
Lecture title	Wood skills	Metal skills	Electronic skills
Safety, tools, equipment	G1	G2	G3
	G2	G3	G1
	G3	G1	G2
Basic electronics and circuits	G1	G2	G3
	G2	G3	G1
	G3	G1	G2
Basic sensors	G1	G2	G3
	G2	G3	G1
	G3	G1	G2
Basic motor and their control	Free time to finalize the end product		

The Hands-On sessions provided ample of time for basic demonstration, practice and troubleshooting and problem solving. The personal guidance was given for the participants. The three hours session was also divided into a three hourly slots. The participants were divided into three smaller homogeneous groups and were rotated every hour through three stations for better involvement. This allowed them to apply what they had learned in each lecture and demonstration to complete their mini project. The participants completed their mini project based on the skill learned with the assistant of the tutors since it was the first time many of them were exposed to these technical aspects.

C. *Innovation Ideas (I)*

At the end of workshop, MAST came out with some innovative ideas. After data collection by unbiased innovators of the university, the potential projects were selected. Analyzing these (I) is an important step to determine the appropriate way to develop it. Some of the innovative ideas could be developed into prototypes within available facilities. Meanwhile, big scale projects required research grants for development.

D. *Project and Prototypes (P)*

Developing of the selected innovative project took place in the creative workstation. The environment was suitable for more experiments and modification. The engineering consultation provided in formal and informal ways to support the development of each prototype. The projects which required special skills and facilities were planned to be executed in engineering school workshop.

IV. ASSESMENT

For evaluation of the cognitive outcome of the workshop, a pre and post training questionnaires were distributed to the participants to analyze the cognitive outcome of workshop. The questionnaire consist of 10 multiple choice questions. These questions are based on the theoretical and practical aspects of the workshop learning objectives.

Meanwhile the practical outcome was monitored throughout the four sessions whilst the participants were constructing their final projects.

By the end of the workshop, a feedback form was distributed to each participant to indicate what the participants liked and disliked in the training and its suitability for their needs and expectations.
(I) are also discussed after the workshop, to be developed and modified in preparation for (P).

V. RESULTS

A. *Creative work station:*

The current workstation has a surface working area of 12 meter^2. It is suitable for most of the heavy and light hardware projects. It is well ventilated and has proper lighting and storage areas. The basic tools are available to work on the convenient materials which is mostly used by any casual user. It is planned to be upgraded in the future in responds to the need and shortage which may be encountered during large scale of practice. Materials and consumables were provided such as: cartoon boxes, light wood, foam, plastics, aluminium sheets and electric circuits.

B. *Learning: Outcome of 1st Mechatronic workshop*

1) *MAST distribution :*

 ISBN: 978-1-61208-167-0

It is found that there is almost an equal interest between the males and females, above and below age of 40 (Table 2)

TABLE II. AGE AND SEX OF THE PARTICIPANTS IN MECHATRONICWORKSHOP

Total	Age Distribution				Sex Distribution	
	>20	>30	>40	>50	Male	Female
23	8	3	4	8	11	12

A total 23 applicants out of 39 enrolled applicants managed to accomplish the two days training program due to concurrent events on these couple of days (Table 3).

TABLE III.DIFFERENT CATEGORIES OF THE PARTICIPANTS IN MECHATRONIC WORKSHOP

Category	n	Notes
Academics/Lecturers	15	4 lecturers attended 1 day training
Non academic MLT Staff	3	A concurrent meeting for administrators held on same dates
Management/Admin	1	
Students	4	6 students could not join the training because of other commitments

2) Practical outcome:

A total seven final models has been accomplished after 4 hands on sessions. It is actually 80% of the expected projects. Those who accomplished the practical projects were excited and showed great enthusiasm. The project is to develop a wooden box with metal bar attached which houses a developed electronic circuit that will be activated by light to on a small fan.

3) Cognitive outcome:

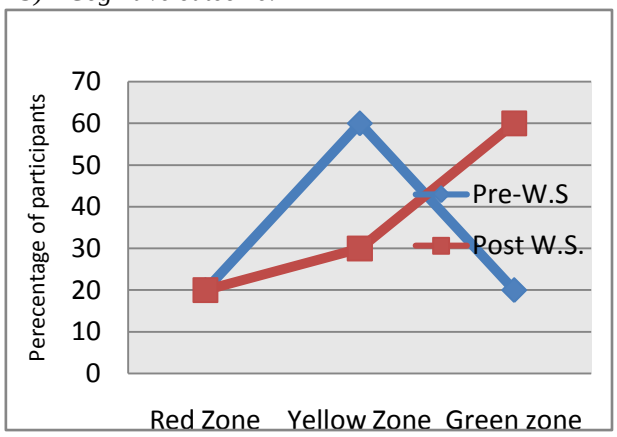


Figure 2. Comparison of the cognitive outcome of the Mechatronic workshop: post W.S (post-workshop) assessment to pre W.S. (pre-workshop) assessment

The results of pre and post workshop test showed a significant improvement of participant's technical awareness. As a standard for evaluation, we considered the score of < 30% as red zone, less than 60% as yellow zone. . The cognitive outcome of post mechatronic workshop is shown in ' Fig. 2"

While more than 60% as a green zone. In the post workshop assessment, the percentage of participants who scored in the green zone was 60%, in comparison to 20% in pre-workshop assessment. Meanwhile red and yellow zones significantly narrowed from 20% and 60% during pre-workshop into 10% and 30% during the post workshop respectively.

4) The participants concern of the learning outcome

All the participants (n=18) were provided with an evaluation form (feedback) at the end of the two days' workshop. The feedback consist of four structured questions to rate the different aspects of the workshop. The 5th question was an open question to express their opinion. Presented here are the results of the evaluation for each question.

The first question of feedback was to survey the overall impression regarding the overall workshop . As shown in "Fig. 3 " , the participants were impressed with exposure to this technical skills. None of the participants provided negative ratings like poor or fair.

Figure 3. Overall impression of Medical School delegates on Mechatronic training

The next questions were to evaluate the program structure and suitability for its objectives are summarized in table 4 . It is rated excellent and good by the majority. As they came to understand how to design a project and to translate the idea to action. In the response to cognitive and new practical skills outcome , most of the participants rated very good or excellent . Most of them were exposed for the first time to such skills . The participants found the new skills are relevant to professional practice and could be applicable in future projects , as shown in " Fig. 4 "

 ISBN: 978-1-61208-167-0

The 5th question gave the participants an opportunity to state what they liked and disliked about the workshop and what improvement to be made in the next training session that they would like to suggest. Varieties of positive comments about the experience of the 2-days training program were recorded.

TABLE IV. THE PARTICIPANTS FEEDBACK REGARDING MECHATRONIC WORKSHOP OUTCOME

Parameter score (m)	Excellent (4)	Good (3)	Fair (2)	Poor (1)	Score mn=72
Program Structure	11	7	0	0	65/72 (90.3 %)
New Knowledge	12	6	0	0	64/72 (88.9 %0
Relevance and Impact	11	7	0	0	65/72 (90.3 %)

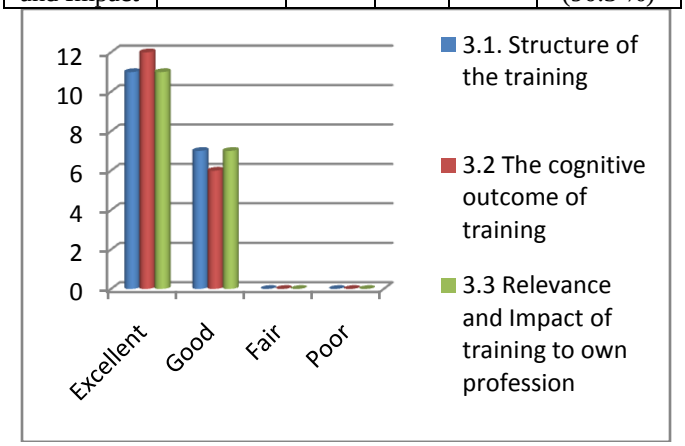


Figure 4. Feed back of participants regarding the contents of the training program

Majority of the participants liked the brain storming atmosphere during the workshop. They expressed that (L) provoked their interest to think out of the box in different aspects . 50 % of them liked the hands-on sessions. 38.8 % was impressed with the lectures, methods of training, new learned skills and knowledge. 11.1 % of participants admired the innovative and interesting training contents.. Academicians were motivated and energized by these technical training. The favorable comments ranged from "The final product is encouraging and comprehending, the training gives clear understanding of the principles, it is challenging, it is great, best workshop attended so far etc." to "it is well organized and properly prepared".

Meanwhile 22.2 % provided negative feedback regarding the shortage of allocated time to finalize the end-product. Another participant expressed about whole day session, especially in post-lunch session . Others pointed out about some shortage of tools , which had been shared .

Many constructive suggestions for the next training sessions came out. Few participant suggested to extend the training period , " from to 3-7 days in a camp like status" . 4/18 suggested to "introducing more medical applications in the program".

2/18 were interested in joining the "advanced level of technical learning in the program". Another participant suggested " to hold the program frequently to give chance for a more to participate". In view of the groups distribution , one student suggested that "each working group to include academics and students together".

C. *Innovative Ideas:*

Out of twenty three participants, seven innovative ideas were laid out as a result of brain storming and exposure to new technical skills. All the ideas came from the lecturers who attended the workshop. Immediately after the workshop, four of these ideas had further study and consultation. (I) underwent modifications and adaptation to transform it into action.

D. *Projects and Prototypes:*

The designated four (I) had the challenge to participate in PEREKA 2011 (Research and innovation competition) at the university level. The four projects had been transformed into protypes using the knowledge, expertise outcome of the mechatronic workshop. As a fruit of (L) , expertise consultation was accessible. These projects came out of academic category (n=15) (26.7 %) over period of three months. Most of the innovators used Creative workstation facilities to accomplish their projects, namely as follows:

1) "Scrubbing and sterilization protection sensors (SPS)".

2) "Teaching aid anatomical model for neural motor pathway in the upper limb"

3) "Amplifier for fetal heart sound for clinical purposes in remote areas and for teaching demonstration." This project has applied for a grant for further development.

4) "CLIP Methodolgy"

Three projects were awarded gold, silver and bronze medals in an PEREKA 2011 (University Innovation competition) on 7th June 2011.

VI. DISCUSSION

Most of changes in the medical practices evolve around science-based health innovation . CLIP methodology aims to encourage the innovation in health services ,which is the synergism of innovative professional products and academic tools ' Fig. 5 "

The professional innovated products vary across a spectrum of sophistication, from vaccines, pharmaceuticals, IT (Information Technology) applications, medical devices to some plant medicines [27] [28]. While in the academic institutes of Medicine, the innovation inputs lead to innovated academic tools as researches , teaching aids and teaching methods [7].

It is an opportunity for lecturers (A) and other MAST categories to go for continuous professional development

 ISBN: 978-1-61208-167-0

[36]. It sharpens the saw of different intelligence of the participants and releases the stress, anxiety and encourages the spirit of learning [35].

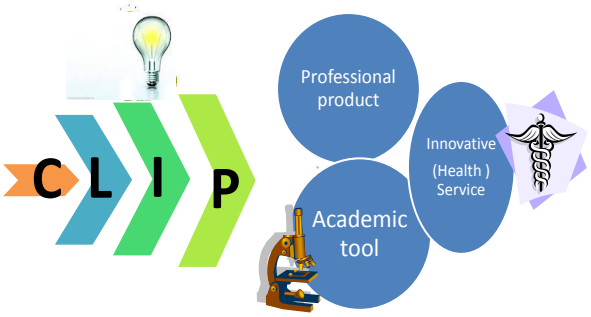


Figure 5. CLIP methodology enhances innovative health services. (C) Creative work station is working venue, (L) learning supplementary skills, (I) Innovative ideas, (P) Prototype production.

Initially , ten students (S) were eager to join the training , but only four of them attended the (L) session due to concurrent school activity. These participants showed enthusiasm and got the highest scores among MAST in post workshop assessment. They expressed their wishes to attend the advanced mechatronic workshop.

When (A) and (S) are involved in the (L), new innovative ways in teaching and learning process develop. Accordingly, it becomes most effective than the dry lectures and boring text books [32]. Exposure to these parallel skills will challenge students to think critically, communicate lucidly, and to think out the box to understand their own curriculum [14][33].

Some academicians (A) had shown some uneasiness on the concept of CLIP as (C) is hardware workshop and (L) content as {mechanical and electronic skills} does not fit with the medical curriculum. Others were afraid to enter unknown area of technical knowledge since most of them are engrained as medical personnel.

Meanwhile many academicians encouraged this new challenging concept. In this preceding trial, 15 academicians (A) had participated in first learning session (L). The post workshop feedback revealed the participants' appreciation to such non-medical skills. Some of the lecturers found it is so interesting to use electronic principles and experiments to teach complicated medical concepts in physiology.

(C) Provides not only the venue, tools and materials for prototype production, but also the consultation services. These facilities had been used effectively to execute some projects. A plan is done to maintain and upgrade (C) to fit different styles of innovations.

A list of (L) skills were chosen to meet the MAST needs and to develop their talents [13][38]. From post workshop feedback, it is clear that longer time for the workshop and additional tools in hands-on session will improve (L) outcome in term mastering the new skills .

The School of Engineering started to apply the structure of (L) ,as a preliminary course for all first year students.

As the trainings are run by collaboration with the expert people in their professions, it stimulate personal and professional growth [31]. The collaboration is a healthy environment that allows the exchange of experiences between engineers, doctors; artists, musicians, architecture and IT people will enrich the university environment known as hybrid learning environment [33].

As a preceding work, other trainings will be arranged in collaboration with experts in the university to enhance MAST` multiple intelligence as: medical photography, fine arts, music, in addition to the advanced mechatronic workshop [29][30]. In this advanced workshop the applications of technology in daily used medical equipments will be highlighted in depth e.g. Fiber optics, endoscopy, manipulating and acquiring medical data. Micro-controller programming which is used in robotic arms will also be covered [39][40]. Such workshops may be a seed for a proposed Bio-Medical Engineering program in the university.

Most of the developed (P) are teaching aids, while the professional (I) had not been translated into action as yet [34]. Through CLIP, synergizing medical basic sciences, clinical medicine with learning (L): non-medical skills in technology is expected to lead to more innovation in health services as shown as in "Fig.6".

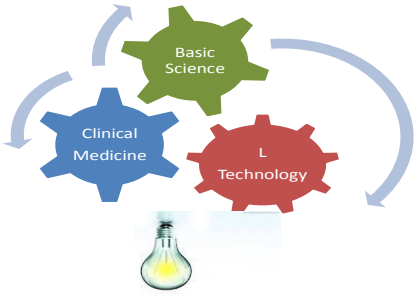


Figure 6. CLIP methodology: Synergism of (L) : technology, basic sciences, and clinical medicine, will lead to more innovation in health services

One intrinsic weakness in this study is limited number of participants MAST. Accordingly further work is required to show statically the impact of CLIP on both academic and professional fields. Persistence of different modalities of training and participation of more candidates of MAST will clarify the effectiveness of this innovative methodology [26]. Other limitation was the limited number of some tools in relation to the number of participants in (L). The present (C) is relatively small and upgrading is required to be able to accommodate transformation more (I) into (P) products. "Innovation is like riding a bicycle; you must keep pedalling or you coast. And, the only way you coast is downhill" [15]

VII. CONCLUSION

CLIP methodology provides the venue, facilities and consultations for the members of the institute to execute

 ISBN: 978-1-61208-167-0

their own innovations. It also exposes them to parallel skills which sharpen their talents. This paradigm gives them the chance to translate Knowledge to Action. In the last 10 months of the activity, brain storming towards innovation had been provoked in the school of medicine. Some innovated projects had been completed and awarded at university level. As a preceding work, a variety of trainings programs such as medical photography will be arranged in the future. CLIP Methodology is potentially suitable for the schools with comparable infra-structure and problems e.g. biology and geology institutes .It is hoped to be an effective methodology in enhancing creativity and innovations.

REFERENCES

[1] Z. J. Acs, & D. B. Audretsch, Innovation and technological change. In Z. J. Acs & D. B. Audretsch (Eds.), Handbook of entrepreneurship research. Boston: Kluwer Academic Publishers, 2003. [1] B. E. Mathews and R. L. Bailey, "A course in creative problem solving,"

[2] Panthalookaran, V.; , "'Hour of Creativity': An agenda to foster creativity and innovation in the students of engineering," *Global Engineering Education Conference (EDUCON), 2011 IEEE* , vol., no., pp.612-617, 4-6 April 2011. doi: 10.1109/EDUCON.2011.5773201

[3] IEEE Transactions on Education, Vol. E-8, No. 4, 1965, pp.85-90.

[4] A. Warsame, P.O. Biney and J. O. Morgan, "Innovations in teaching creative engineering at the freshman level," 1995, Proceedings of ASEE/IEEE Frontiers in Education Conference, Session 2C4, pp. 2C4.21-2c424.

[5] L. G. Richards, "Stimulating creativity: teaching engineers to be innovators," Proceedings of 28 th ASEE/IEEE Frontiers in Education Conference, Session S1D, 1998, pp. 1034-1039.

[6] Thakur R, et al, Innovation in healthcare: Issues and future trends, J Bus Res (2011), doi:10.1016/j. jbusres.2011.02.022

[7] http://en.wikipedia.org/wiki/Research [accessed: Sept. 10, 2011]

[8] http://www.increasebrainpower.com/innovation-brainstorming.html [accessed: Sept. 10, 2011]

[9] Incremental innovations in India @ http://www.articlesbase.com/intellectual-propertyarticles/incrementalinnovations-in-india-4138579.html [accessed: Sept. 10, 2011]

[10] http://en.wikipedia.org/wiki/Innovation [accessed: Sept. 10, 2011]

[11] http://www.informationarchtected,com/surfaces/innovation-management [accessed: Sept. 10, 2011]

[12] http://ezinearticles.com/?Types-of-Innovation&id=38384 [accessed: Sept. 10, 2011]

[13] http://www.eiu.com/site_info.asp?info_name=doctor_innovation&page=noads&rf=0http://www.intouchhealth.com/ [accessed: Sept. 10, 2011]

[14] http://www.willatworklearning.com/2006/05/people_remember.html [accessed: Sept. 10, 2011]

[15] G. C. Nicholson, "Keeping innovation alive". Res.Technol. Manage., vol. 41 (3), pp. 34—40, 1998.

[16] G. T. Lumpkin, G. E. Hills, & R. C. Shrader, Opportunity recognition, in entrepreneurship: The way ahead, New York: Routledge, 2003.

[17] G. T. Lumpkin and B. B. Lichtenstein, "The role of organizational learning in the opportunity-recognition process". Entrepreneurship: Theory & Practice, vol. 29(4), pp. 451-472 , 2005.

[18] C. D. Aubrey, Performance Management: Changing Behavior ThatDrives Organizational Effectiveness, 4th ed. Performance Management Publications, 2006.

[19] Love, J.H., et al., Openness, knowledge, innovation and growth in UK business services. Res. Policy (2011), doi:10.1016/j.respol.2011.05.016

[20] D. Cyr, "How to solve almost anything". Attache, pp. 46-48. November 1997.

[21] Senge, Peter M. , *The Fiph Discipline: The **Art** & Practice of The Learning Organization*, New York: Doubleday. New York, 1990.

[22] Leonard, D. and Swap, W., "When sparks fly: Igniting creativity in groups", Workforce, Vol. 78, No. 10, 1999, pp. 87-89.

[23] Ekwaro-Osire, S. and Orono, P., "Pan-mentoring in creative engineering design - the coordination of individual and team creativity," in Proceedings of 2005 Frontiers in Education Conference, Indianapolis, Indiana, 2005, pp. S2J-7-S2J-11.

[24] Carkett, R., ""He's different, he's got 'Star Trek' vision': Supporting the expertise of conceptual design engineers", Design Studies, Vol. 25, No. 5, 2004, pp. 459-475.

[25] Dorst, K. and Cross, N., "Creativity in the design process: Co-evolution of problem-solution", Design Studies, Vol. 22, No. 5, 2001, pp. 425-437.

[26] Hargadon, A., How breakthroughs happen: The surprising truth about how companies innovate, Boston, Havard Business School Press, 2003.

[27] S. Al-Bader, A. S. Daar, and P. a Singer, "Science-based health innovation in Ghana: health entrepreneurs point the way to a new development path.," *BMC international health and human rights*, vol. 10 Suppl 1, no. 1, p. S2, Jan. 2010.

[28] J.M. Devoge , E.J. Bass, R.M. Sledd , S.M. Borowitz , and L. Waggoner-Fountain, "Collaborating With Physicians to Redesign A Sign-out Tool an Iterative, Multifaceted Approach With Users – Even Busy Ones – can Yield a Satisfying and Efficient Product." Ergon Des. 1;17(1):20-28, Jan. 2009. doi: 10.1518/106480409X415170

[29] Gardner, Howard. Multiple Intelligences: The Theory in Practice. New York: Basic, 1993.

[30] Armstrong, Thomas. Multiple Intelligences in the Classroom 3rd ed. Alexandria, VA: Association for Supervision and Curriculum Development, 2009.

[31] K. K. Schultz and M. L. Brackbill, "Teaching electrocardiogram basics using dance and movement.," *American journal of pharmaceutical education*, vol. 73, no. 4, p. 70, Jul. 2009.

[32] Armstrong, Thomas. In Their Own Way: Discovering and Encouraging Your Child's Personal Learning Style, New York: Tarcher/Putnam, 1987.

[33] R. A. Blouin et al., "AACP Curricular Change Summit Supplement Roles of Innovation in Education Delivery," *American Journal of Pharmaceutical Education*, vol. 73, no. 8, 2009.

[34] J. P. Lane and J. L. Flagg, "Translating three states of knowledge--discovery, invention, and innovation.," *Implementation science : IS*, vol. 5, p. 9, Jan. 2010.

[35] J. Shapiro et al., "Relationship of creative projects in anatomy to medical student professionalism, test performance and stress: an exploratory study.," *BMC medical education*, vol. 9, p. 65, Jan. 2009.

[36] http://en.wikipedia.org/wiki/Teacher_education/ [accessed: Oct. 10, 2011]

[37] L. K. J. Baartman and E. de Bruijn, "Integrating knowledge, skills and attitudes: Conceptualising learning processes towards vocational competence," *Educational Research Review*, vol. 6, no. 2, pp. 125-134, Jan. 2011.

[38] http://rapidbi.com/management/six-big-problems-with-training-and-development/ [accessed: Oct. 10, 2011]

[39] G.Rizzoni.2007. Principles & Applications of Electrical Engineering (5thEdition). McGraw Hill

[40] Hughess, E. 2002.*Electrical & Electronic technology*.(8th Edition). Prentice Hall

 ISBN: 978-1-61208-167-0

Positive Psychology Centered Online Studies

Alexander Voiskounsky
Psychology Department
Moscow Lomonosov State University
Moscow, Russia
e-mail: vaemsu@gmail.com

***Abstract*—Flow (optimal) experience is a well-elaborated unit in positive psychology. This unit fits well many activities mediated by information and communication technologies. The paper covers the studies in the mentioned above area, which were carried out by Russian psychologists during the last decade. The empirical studies include the analysis of flow experience in online gaming and of its role in computer hackers' motivation; theoretical proposals refer to optimal experience in the web and software usability research and an appeal to differentiate Internet addiction and flow experience. The paper states that theoretical and applied work based on the optimal experience methodology is a promising perspective for online research.**

Keywords-flow; psychology; optimal experience; information technologies; online gaming; hackers; usability; addiction.

I. INTRODUCTION

From a psychological standpoint, virtual space is a human centric space, intermediate between a human being and others, between an inner and outer knowledge, between self and multiple selves. Being mediated by computers (or mobile vehicles) and the Internet, virtual space is characterized by its own inherent metrics (dimensions): neither *X/Y/Z* Cartesian coordinates with *up/down, left/right, and forward/backward* directions of movement, nor Einsteinian *spacetime* are easily and plausibly applicable in the virtual space. At the same time, the mentioned above genuinely scholastic dimensions are the ones which everyone who enters a technology-rich working environment expects to face.

Virtual spaces represent comfortable environments for almost everyone, especially for new generations, often called "digital generations", or "digital natives" [7], starting with an elementary school age or even earlier; virtual spaces are not any more the privilege of students and lecturers in computer science, engineers and all those who are engaged in academic work. To enter a virtual space, one need not leave either office or home, or both. Consequently, the dimensions describing virtual spaces need not be entirely scholastic any more; on the contrary, these dimensions are becoming more and more close to the ones which proved its usefulness in descriptions of contemporary industrial, office or home environments.

For example, the most part of the well-known dimensions of the habitual 3D space, such as for example an *office* area as opposed to a *living* and *sleeping* area opposed to a *recreational* area opposed to a *shopping* area opposed to an *interaction* area opposed to a *private* area (all the oppositions are not overwhelmingly strict, of course) have definite parallels in the virtual space. Indeed, the virtual space dimensions are in a way similar to what people are apt to do both in the habitual (physical, or real-life) space and in the virtual space, namely: *work, interact and make relations, play and recreate, learn and explore, go shopping, negotiate and gamble, save privacy*, etc.

Many people find these dimensions easy to follow; but there are those who believe the organization of the virtual space dimensions is far from being optimal. Very likely, this distinction reflects the life-style differences and corresponds to diverse human experiences. Anyway, what is known about attempts to optimize these dimensions? Optimization is dependent on numerous factors, such as economics, education, logistics, profession, personality traits, etc.; this Section however deals with psychology, and thus we will discuss psychological parameters of optimal forms of virtual space dimensions. In particular, some emotion-related dimensions such as *enjoyment* or *pleasure* are no less universal than the Cartesian ones, whenever people tend to make assessments of unknown spaces. That means, we are going to discuss in detail how well a psychological theory of optimal experience fits the virtual space.

In the paper, we review the current directions of studies and give a perspective of optimal experience related academic and practical work within the virtual environments. The review is centered on the studies which we have done during the last decade.

II. OPTIMAL (FLOW) EXPERIENCE: A BRIEF DESCRIPTION

Principles and methods of positive psychology [10] are widely used in the growing research field of human behavior in virtual environments. The most promising perspectives in positive psychology seem to be the Self-Determination theory developed by Deci and Ryan [2] and the optimal (flow) experience theory developed by Csikszentmihalyi [1]. The methods belonging to the positive

 ISBN: 978-1-61208-167-0

psychology and particularly to flow experience relate intimately to the patterns of human behavior in virtual environments, such as interaction, learning and exploration, problem solving, gameplaying, shopping, etc. The history and the current status of the positive psychology studies referring to the virtual environments was recently reviewed [5][15]. The discussion in the review papers is restricted to the studies related to optimal experience theory, due to the fact that the studies related to virtual environments and influenced by the Self-Determination theory are not numerous: rather few publications deal with e-learning [9] or online gaming [8].

The feeling of finding oneself – mentally – in the midst of a torrent of liquid (especially running water) is not totally alien to psychological theory, taking into consideration traditional terms such as stream of consciousness or flight of ideas. The notion of flow has been introduced by Mihaly Csikszentmihalyi who interviewed hundreds of people (among them painters, athletes, physicians, dancers, scientists and many others), asking them to describe their feelings while carrying out their professional or hobby work. Each time when respondents expressed a really deep devotion to this or that sort of work (not necessarily pleasing and playful, often, on the contrary, hard and risky), their reports contained a common element, or metaphoric description of a sensational experience for which Csikszentmihalyi [1] could have hardly choose a name other than **flow**. Indeed, almost every respondent mentioned "flowing from one moment to the next, in which he is in control of his actions, and in which there is a little distinction between self and environment, between stimulus and response, or between past, present, and future" [1, p. 36]. Flow (both deep and rare flow events, and habitual *microflow* events) was reported to happen irrespectively of the type of the work: be it spiritual or mundane, creative or routine, unique or known to almost everyone, individual or team-work, rarely or regularly performed.

Csikszentmihalyi and his followers found that respondents never report of flow happening (and it is indeed a sort of a happening!) when they feel relaxed: on the contrary, to experience flow they need to be genuinely and deeply involved in the preferred work. People report they experience flow when they perform their work to the utmost and get a positive result: Csikszentmihalyi and other scholars would call it a **peak** performance. The necessity to achieve some success, which may be in fact quite moderate, explains the need to acquire, prior to experiencing flow, some competence, not necessarily very high, in performing the work. Respondents always describe flow as an **enjoyment**, which is characteristic of flow; in the lifespan they tend to remember such happenings quite well, sometimes decades after the experience itself. Clearly, within the positive psychology paradigm this sort of experience is called **optimal experience**.

Positive psychology is a universal psychological discipline; when applied in practice, however, it works probably best as a source of motivational theories: this refers both to the Self-Determination theory and to the theory of optimal experience. Neither is an exclusively motivational theory; both are used in practice as theories of **intrinsic** motivation. Various types of such motivation have been widely exploited while designing computer/Internet applications, as well as in the practice of self-regulation, management, education, etc. An intrinsically motivated process is self-rewarding, while its results might be (at least partly) irrelevant. In the optimal experience context a process, or a sequence of intermediate goal-directed efforts performed in order to achieve the desired result is quite often reported to be much more pleasing and self-rewarding than the result itself, when and if it is gained. Optimal experience has been called [1] **autotelic** (from Greek: self + goal), it means that the goal of doing some work is just the act of doing it, regardless of whether external rewards will follow. While an intrinsically motivated work brings *enjoyment*, Csikszentmihalyi refers an **exotelic** work and extrinsic motivation to *pleasure* which is a somewhat passive and relaxing feeling, compared to enjoyment, at least in the optimal experience context. No doubt, the achievement of pleasure, and respectively the extrinsic motivation (in the form of money bonuses, lovely sex partners or increases of power) have always been an incomparably strong stimulus of diverse human activities. Positive psychology suggests, nevertheless, that neither intrinsic stimuli nor enjoyment should be underestimated; successive managers try to combine the two types of motivation to stimulate the employees.

III. FLOW CHARACTERISTICS AND DIRECTIONS OF STUDY OF OPTIMAL EXPERIENCE IN THE VIRTUAL ENVIRONMENTS

Theoretically and empirically, Csikszentmihalyi [1] selected the following major – or the most common – **characteristics of flow**:

- ✓ *clear and distinct objectives;*
- ✓ *temporary loss of self-consciousness;*
- ✓ *distorted sense of time;*
- ✓ *actions merging with awareness;*
- ✓ *immediate feedback;*
- ✓ *high concentration on the task;*
- ✓ *high level of control over the task;*
- ✓ *balance (precise matching) between the available skills and the task challenges;*
- ✓ *full satisfaction, while doing work, which is worth doing for its own sake*.

This set of characteristics proved to be applicable in diverse environments [1], including the use of information and communication technologies: the relevant studies took a start in early 1990s [8][9]. In the virtual environments, which are the main point of interest in the current paper, the most frequently suggested characteristics, additional to the

 ISBN: 978-1-61208-167-0

mentioned above, include (see the appropriate references at [15])

- ✓ *presence* ("mediated perception of an environment" or "being there – in a somewhat different place", as well as "possibly sharing this place with other people", and finally "immersion"), and
- ✓ *interactivity* (distinguishes new – responsive – media from traditional media).

Each of the two additional characteristics has been intensely applied and investigated in empirical studies and often proved its usefulness. One can formulate that characteristics such as *presence* and *interactivity* are among the most needed within the virtual environments, partly due to the fact that these characteristics are inherently related with the information and communication technologies. In a book chapter written with Shernoff, the originator of flow related studies accepts recently proposed characteristics such as "presence" or "being there" as well as "immersion", and makes it evident that "flow theory has been the natural theoretical base for exploring the implications of learning through immersion in … virtual learning environments since the emotional composition of these experiences resemble the flow state and precipitate a deeper engagement with learning" [11, p. 141].

The practice of the use of new information and communication technologies represents a variety of areas to study flow experience. From the very beginning [4][13] the studies of optimal experience in the virtual environments referred to 'computer-mediated communication', 'online, 'computer based instruction', 'Web use', 'human-computer interaction', 'computer-mediated environments' related to marketing, 'information and communication technology use', 'activities involving information technology,' 'human-computer interaction' or 'Internet/Web use'.

When classified, the major areas of the use of the flow experience methodology in virtual environments are the following [15]:

- Online marketing and shopping,
- E-learning/teaching,
- Cyber-recreation (most often, online, computervideo gaming),
- Virtual interaction.

These research areas do not differ from traditional directions of studies which are being carried out in psychology of cyberspace. Additionally, less elaborated research areas include [15], and are partly discussed in the paper:

- Virtual psychological rehabilitation, such as immersive systems of virtual reality;
- Illicit penetrations into the virtual space environments and computer security regulations;
- Usability testing, measurement of a web-site attraction and friendliness, adaptation of web sources to target populations.

In this paper, we are not going to discuss all the mentioned above directions of the studies and applied work. Instead, we will limit with the description of the studies held in the field by the author and his colleagues and students during the last decade. In spite of the fact that the most part of these studies have been held in Russia within the last decade (see the appropriate descriptions and links to the references in Section IV), we believe that the results are cross-cultural by their nature.

IV. EMPIRICAL STUDIES OF FLOW EXPERIENCE IN THE VIRTUAL ENVIRONMENTS

A. Theoretical Proposals

In this Section we are going to discuss two proposals which lack empirical support to be done by the author. First, we will discuss the need to connect the optimal experience methodology with applied work traditionally done within the projects targeted at the development of new software products, namely the work which is known as usability testing. The second point to be discussed in this Section deals with a possible though non-likely (as it is stressed further in the Section) correspondence between the flow experience and the Internet addiction disorder (also known under diverse names, such as for example Internet overuse, Problematic Internet use, Pathological Internet use, Internet abuse, or Compulsive Internet use, etc.).

Current methodology of usability testing should rely on longitudinal research and on field studies. When a longitudinal fieldwork usability project is being carried out, it is preferable to investigate the users' intrinsic motivation which has the highest prognostic value. On the contrary, traditional lab settings are optimal to investigate extrinsic motivation; since intrinsic motivation is unlikely to be revealed while carrying out a traditional type of usability-related work, the importance of motivational research has been largely underestimated.

A promising approach toward better understanding of specifics of intrinsic motivation is a flow paradigm [12]. Good match between the software users' needs and skills, on the one hand, and the inner structure of programming tools, on the other hand, means that the choice of the "next step" while fulfilling the task needs to match the (possibly) increased skills.

The usability practice faces the problem of checking whether customers keep experiencing flow in a long perspective. Usability engineers need an advanced specialized methodology to apply in field work. In the information and communication technologies field, special software can be developed to assist the participants in reporting particular characteristics of flow, or alternatively, the absence of these characteristics. This methodology seems to be extremely promising for carrying out

 ISBN: 978-1-61208-167-0

longitudinal fieldwork, including software users' motivational research [12]. The alternative models, which are widely used to evaluate usability of software products, only rarely rely on a registration of indisputably fundamental type of human needs such as intrinsic motivation. Due to this fact the suggested method of usability studies can be used either parallel to other methods, or alone. Even in the latter case, one may expect the results to be exceptionally reliable. When an applied usability study involves longitudinal evaluation, the methodology of optimal experience is really promising, irrespectively of how many alternative methods are being used.

The second point to mention refers to the Internet addiction disorder in relation to the optimal experience. Indeed, there is a growing body of evidences – mostly referring to gameplay activities – in which flow experience is correlated with the Internet addictive behavior (see appropriate references in [14][15]). At first glance, there is indeed some reason to correlate the two types of experience, related to addiction and to flow, since repetitive behaviors characterize both types of activities. Nevertheless, flow is an example of a *positive optimal* experience and it is hardly compatible with addiction or psychological dependency, typical for a *negative* kind of experience [14][15].

Repetitive behavior referring to a negative type of activities, such as intakes of drugs, passive leisure activities or abuse of technological artifacts, including computers and computer games, is derived from the so-called *mimetic flow* [6], i.e. non-genuine flow experience. Alhough it may result in replications of a familiar behavior, the psychological nature of mimetic flow is dissimilar with the nature of optimal experience, as it was described earlier in this paper. Thus, the Internet addiction and flow experienced while using the Internet differ in their inner psychological structure; this theoretical statement is apt to get an empirical support, i.e. evidences that flow experience correlates negatively to Internet addiction disorder.

B. Flow in Computer Hackers' Activities

The study of optimal experience patterns in computer hackers' behavior was done as an online study [19]. The study rests on an idea that in order to experience flow a balance is needed between the level of software use (not specifically hacking) skills and the level of challenges (or task choices) in hacking.

A hacker's development might be presented in the following way. The relationship between hackers' experience and flow is complicated. Flow does not linearly increase with the increase of the hackers' competence. Periods of flow experience turn to periods of flow crisis and then to periods of flow renovation. In the post-experimental interviews the role of task choice in experiencing flow while hacking is revealed. The step-by-step task choice often leads to close matching of task challenges and skills, marking flow experience.

An inexperienced hacker (a beginner) might find a matching combination of challenges and skills and start to experience flow. The flow motivation is strong, and the beginner feels comfortable. A hacker might stay at this stage for years. To stay at a beginner's stage means that neither skills nor challenges develop in a significant way.

A beginner hacker might progress in at least **three** ways. The *first* is a step-by-step progress both in challenges and skills which keep matching at every developmental stage. The progressing hacker keeps experiencing flow all the time. The cases of fine skills/challenges correspondence at every stage are probably infrequent. *Second*, a hacker gains new skills and lacks the correspondence of new skills to non-updated challenges. Or, *third*, a hacker takes high challenges and finds he/she lacks non-updated skills. These two ways of a hacker's progress result in periodical dropouts of the flow range, and the hacker periodically stops experiencing flow. Constant matching of skills and challenges and non-interrupted flow seems to be a hard way of progressing.

If an inexperienced hacker increases challenges, he/she turns to become a wannabe hacker, at least for a certain stage of his/her development. A hacker might stay at this stage long enough, trying to acquire prestigious goals and never acquiring them. A wannabe hacker's rewards might lie in the social life: he/she might boast and get some social prestige. To renovate the flow experience, a wannabe hacker might either lessen challenges and to become an averagely competent hacker setting moderate challenges, or update skills to become a highly qualified hacker with challenges matching the available skills.

If an inexperienced hacker increases skills, he/she looses the fine matching of challenges and skills. His/her skills overrun his/her challenges until the challenges are not updated. When updated, the flow experience might come back at a higher level of skills/challenges correspondence. Former hackers often turn into computer security officers; that means they loose motivation to pose high challenges in hacking.

The study, which was briefly described in this Section, shows first, that the hackers' motivation includes experiencing flow, and second, that the hackers' motivational development is strongly dependent on flow experience while this development includes longer or shorter periods of interruption and/or dropout periods when a hacker ceases to experience optimal forms of motivation.

C. Flow Experience in Online Gaming: A Cross-Cultural Study

Online gaming is among the most popular patterns of use of online services. One of the potential reasons of this popularity is that the gamers are fond of playing video and online games due to the fact that they experience flow while they perform gameplay. This reason was investigated in a number of studies, including a cross-cultural Russia/France/USA/China (the latter study is still in

 ISBN: 978-1-61208-167-0

progress) study with an identical methodology. In all the studies the methodology is the same: first we worked out a comprehensive questionnaire in Russian, and after it proved to provide reasonable results, this flow-related questionnaire was adapted to be used in different ethnic communities of gamers [3][16][17][18].

Each time we performed translations and back translations until the questionnaire proved to be adequate. The methodology included online administered surveys; we did not seek gamers playing identical games since we needed replies from the ethnic communities of gamers who played the whole continuum of online games. Here is some statistics related to the results gained in the series of multiethnic factor analytic studies. The Chinese data are nowadays being studied anew since the gamers' population is steadily growing, and the data need to be re-studied on the current stage of development of the Chinese gamers' population.

Every particular ethnic-specific factor model proved to be reasonably good; moreover, the factor structures seem to be similar in many ways. Most important, each factor model includes a factor such as 'flow experience', mostly as the first factor (except for the French sample). The results are represented in the following table (see Table I).

TABLE I. COMPARISON OF FACTOR MODELS RELATED TO ETHNIC SAMPLES OF ONLINE GAMERS

Ethnic Samples	**Russians** N = 347	**Chinese** N = 133	**French** N = 202	**Americans** N = 287
Factor 1	Flow	Flow	Achievement	Flow
Factor 2	Achievement	Achievement	Interaction/ Cognition	Achievement
Factor 3	Activity/ Passivity	Spend Time	Flow	Cognition
Factor 4	Interaction	Interaction		Self-Control
Factor 5	Thoughtful/ Spontaneous			Interaction
Factor 6	Cognition			

The relevant factor structures cannot be presented and discussed in this paper in detail; confirmatory factor analysis was done in the appropriate empirical studies. The thorough analysis is presented in the papers [3][16][17][18].

The factor structure of the Russian gamers is the most complex (among the samples we investigated) and contains more factors (namely, six) than the factor structures of the other ethnic groups of gamers. All the factor structures are well-established and reasonable; flow is the first factor in almost all the models, with the single exception of the French population of online gamers. The corresponding factor model for the latter sample contains only three factors, which looks minimalistic among the factor structures we examined.

The French sample seems to be even more peculiar since Cognition and Interaction merge into one factor. A likely supposition is that the members of the French sample are fully aware of the fact that interactions with other players provide a chance to exchange and share in-game experience, which is equivalent to performance of cognitive actions in a comfortable manner. Indeed, within the research project we discuss, cognition refers to knowledge useful for making game-related decisions; in an online multiplayer mode of gaming, social perceptive knowledge pertaining to other players is no less valuable than information about the specifics of gameplay structure; exceptionally useful knowledge referring to social perception of partner players can be gained by means of interactions with them. Due to this fact the interactive mode of online gaming is highly appreciated by players. Turning to the factor model characterizing the Chinese sample it is evident that the members of this sample are the only ones (among those investigated) who seem to underestimate processes of collecting knowledge: in the corresponding factor model there is no factor responsible for cognition.

Conclusions such as the mentioned above become evident after the elements of factor models which have been gained in the process of carrying out particular culture-related studies are organized in a table format. This format seems to be helpful: even simple comparison of the existing factor structures can provide hints for new hypotheses. For example, one of the reasons to perform a new mentioned above study within the same project, namely the study dealing with a rapidly growing population of Chinese online gamers, is our intention to find out whether online gamers from China still keep restraining from cognition while gaming.

The main result of the cross-cultural study says that flow experience seems to be an important and significant factor in online gaming and is one of explanations of its ever growing popularity.

V. CONCLUSION

Flow experience, or optimal experience, as introduced by Csikszentmihalyi, takes an important place in the field of the use of information and communication technologies. In a number of theoretical and empirical studies this view got a sufficient confirmation. A portion of significant studies in the field has been done by a team of psychologists from Russia; in the current paper this contribution to the field has been overviewed briefly. A thorough and fairly recent review of the world-wide studies devoted to the use of optimal experience methodology in virtual environments is presented elsewhere [15].

 ISBN: 978-1-61208-167-0

It has been shown that flow experience is a common motivating element for online gamers in four different ethnic groups such as Chinese, Americans, French and Russians. This proves the fact that optimal experience is a basic element in one of the most well-grounded explanations of world-wide attractiveness of playing massive multiplayer online role-playing games, and very likely other types of video and/or console games.

A nonlinear model of the development of computer hackers' flow experience has been worked out: flow depends on the balance between the challenges the hackers take and the skills they possess (skills in competent computer usage, not necessarily in the use of specific software programs for hackers). Flow is shown to develop in a step-like manner which includes "flow crisis gaps", i.e. shorter or longer time periods during which hackers do not report of optimal experience. The flow crisis gap periods are promising for the hackers' dropouts: their eventual transformation into non-hackers, possibly (but not necessarily) into experts in computer security; such a transformation may happen when a hacker becomes preoccupied with mastering his or her skills in computer science and/or software development, and is losing interest in new hacking related challenges. This model shows prospective directions of educational work aimed at reducing the number of computer hackers by transforming them into qualified experts in computer science, including for example the problem area of computer security, or any other problem area which responds to their particular interests.

Flow experience is a complex motivational state and should not be confused with some other psychological states, such as for example a computer/Internet addiction. Flow belongs to entirely positive states which are often easy to be mixed up with creativity, while any addiction is commonly believed to be a psychological state which is desirable to get rid of. It is shown in the current paper that optimal experience is a valuable parameter to be used in performing web/software usability studies.

For the youngest generations – those *born digital* – as well as for the representatives of older generations the cyberspace is becoming more and more habitual, just the space for doing office work, for learning and exploration, for spending leisure time, for interactions, for shopping, for self-presentation, for diverse forms of entertainments (first of all – gaming, viewing videos and movies, uploading photos and videos), etc. It is essential that this virtual space should be comfortable and easy-to-operate, both emotionally and ergonomically. The optimal experience theory developed by Csikszentmihalyi represents an underestimated (in the earliest practice of designing the cyberspace) dimension which might be helpful, as is shown in the current paper, in constructing the virtual spaces which are emotionally comfortable; besides, the current paper shows that this theory provides valuable targeted hints for performing high-quality usability related studies and for applied work aimed to enhance the comfort and effectiveness of using both current and prospective software products.

Thus it means that flow experience should be considered one of the major and most constructive dimensions pertinent for estimating new and old virtual environments and for designing/redisigning these environments. The notion of flow is an entirely human-centric dimension of virtual environments, and from now on this dimension should not be underestimated. Both scholars and producers of new technologies need to work jointly in order to establish reliable standards to ensure that the would-be users of prospective virtual environments are likely to experience flow. This type of theoretical and empirical work, while being of global importance, suggests that it might be performed in a series of cross-cultural projects; the current paper illustrates a particular pattern of such a project.

ACKNOWLEDGMENT

The studies on which the current paper is based have been supported in 2002-2009 by the Russian Foundation of Humanities and by the Russian Foundation for Basic Research.

REFERENCES

[1] M. Csikszentmihalyi, Beyond boredom and anxiety: Experiencing flow in work and play. San-Francisco: Jossey-Bass, 2000 (first published in 1975).

[2] E. Deci and R. Ryan, Intrinsic motivation and self-determination in human behavior. New York: Plenum Press, 1985.

[3] A. Faiola and A.E. Voiskounsky, "Flow experience of MUD players: Investigating multi-user dimension gamers from the USA," Lecture Notes in Computer Science, vol. 4564, 2007, pp. 324-333.

[4] J.A. Ghani and S.P. Deshpande, "Task characteristics and the experience of optimal flow in human-computer interaction," The Journal of Psychology, vol. 128(4), 1994, pp. 381-391.

[5] D.L. Hoffman and T.P. Novak, "Flow online: Lessons learned and future prospects," Journal of Interactive Marketing, vol. 23(1), 2009, pp. 23-34.

[6] F. Massimini and A. Delle Fave, "Individual development in a bio-cultural perspective," in American Psychologist, vol. 55(1), 2000, pp. 24-33.

[7] J. Palfrey and U. Gasser, Born Digital: Understanding the First Generation of Digital Natives. N.Y.: Basic Books, 2008.

[8] A.K. Przybylski, C.S. Rigby, and R.M. Ryan, "A motivational model of videogame engagement," Review of General Psychology, vol. 14, 2010, pp. 154-166.

[9] J.C. Roca and M. Gagne, M. "Understanding e-learning continuance intention in the workplace: A self-determination theory perspective," Computers in Human Behavior, vol. 24(4), 2008, pp. 1585-1604.

[10] M.E.P. Seligman and M. Csikszentmihalyi, "Positive psychology: An introduction," American Psychologist, 55(1), 2000, pp. 5-14.

[11] D.J. Shernoff and M. Csikszentmihalyi, "Flow in schools: Cultivating engaged learners and optimal learning environments," in Handbook of Positive Psychology in Schools, R. Gilman, E. S. Huebner, and M. Furlong, Eds. New York: Routledge, 2009, pp. 131-145.

 ISBN: 978-1-61208-167-0

[12] O.V. Smyslova and A.E. Voiskounsky, "Usability studies: to meet or not to meet intrinsic motivation," PsychNology Journal, Vol. 7, № 3, 2009, pp. 303–324. Retrieved March 14, 2010, from http://www.psychnology.org/File/PNJ7(3)/PSYCHNOLOGY_JOURNAL_7_3_SMYSLOVA.pdf

[13] L.K. Trevino and L. Webster, "Flow in computer-mediated communication," Communication Research, vol. 19(5), 1992, pp. 539-573.

[14] A.E. Voiskounsky, "Two types of repetitive experiences on the Internet," INTERFACE: The Journal of Education, Community and Values, vol. 7, issue 6, 2007. Retrieved February 21, 2011, from http://bcis.pacificu.edu/journal/article.php?id=174

[15] A.E. Voiskounsky, "Flow Experience in Cyberspace: Current Studies and Perspectives," in Psychological Aspects of Cyberspace: Theory, Research, Applications, A. Barak, Ed. N.Y.: Cambridge University Press, 2008, pp. 70-101.

[16] A.E. Voiskounsky, O.V. Mitina, and A.A. Avetisova, "Playing online games: Flow experience," PsychNology Journal, Vol. 2, № 3, 2004, pp. 259–281. Retrieved October 22, 2008, from http://www.psychnology.org/PSYCHNOLOGY_JOURNAL_2_3_VOISKOUNSKY.pdf

[17] A.E. Voiskounsky, O.V. Mitina, and A.A. Avetisova, "Communicative patterns and flow experience of MUD players," Int. J. Advanced Media and Communication, Vol. 1, № 1, 2005, pp. 5-25.

[18] A.E. Voiskounsky, O.V. Mitina, and A.A. Avetisova, "Flow experience and interaction in online gaming: Comparative study of Russian and Chinese MUD players," in Sixth international conference on Cultural Attitudes Towards Technology and Communication 2008, F. Sudweeks, H. Hrachovec, and Ch. Ess, Eds. School of Informational Techology Murdoch University, 2008, pp. 410-421.

[19] A.E. Voiskounsky and O.V. Smyslova, "Flow-based model of computer hackers' motivation," CyberPsychology & Behavior, Vol. 6, № 3, 2003, pp. 171-180.

 ISBN: 978-1-61208-167-0

A Study on the Lost Seeking Devices and Systems for Dementia-Patients

Yung-Ching Chen
The Graduate Institute of Design Science
Tatung University
Taipei, Taiwan (R.O.C.)
e-mail: d9804001@ms.ttu.edu.tw

Cherng-Yee Leung
The Department of Industrial Design
Tatung University
Taipei, Taiwan (R.O.C.)
e-mail: leung@ttu.edu.tw

***Abstract*—This paper utilized a user-centered design approach as the foundation for technology in dementia care in order to improve the quality of telemedicine service. A status-quo analysis and questionnaire survey were conducted to explore the actual needs of the elders in using the lost seeking devices and the problems they encountered. 37 caregivers for people with dementia were surveyed. Through analysis and induction, 3 problems were identified: poor information transmission, low user acceptance, individual material security anxiety. 2-4 improvement proposals are suggested for each problem.**

Keywords-Care-givers; Dementia; e-Health; Safety; Telemedicine services; User-Centered Design; Wandering

I. INTRODUCTION

In the e-era, e-Health is an important issue. Telemedicine service is a means to realize the e-health. In order to have good quality of telemedicine service, the needs and requirements of users are essential elements. Nowadays, some institutes apply this kind of service through lost seeking devices to help find lost people, especially for dementia-patients. However, the users' behaviors and opinions related with these systems and devices are still unknown. Based on the user-centered design (UCD) concept, this paper aimed to find the needs and requirements from users, and expected the results could be used as the foundation for technology in dementia-patient care.

Dementia is now recognized as a global issue of increasing importance, affecting some 24 million people around the world [10]. In Taiwan, the population of demented agedness has been up to a million. People with dementia experience progressive cognitive impairments that typically begin with short term memory problems but can encompass language deficits, difficulties initiating tasks, planning, monitoring and regulating behavior, and visuospatial difficulties [9]. The neuropsychiatric changes in dementia are nearly universal and may result in extremely challenging management problems [14], and heavy burden on caregivers.

However, wandering behavior is the reason that many family members unable to take care the demented elders in the house [2][15]. The most controversial application of technology for dementia care in e-Health and telemedicine services is seen in the application of tracking and surveillance equipment. Proponents of these technologies argue that the technologies reduce caregiver stress and increase individual autonomy since they allow the person with dementia having freedom to move around as they wish [12][13]. Critics however argue that these technologies are an infringement of people's civil liberties and undermine the personhood of the individual [8]. However, assistive technology is a key aspect of improving healthcare. Bjørneby et al. [5] stated that the technologies should: 1) give a feeling of independence to the person; 2) support the person in making choices; 3) have a positive impact on his/her life; 4) support skills maintained or do not emphasize lost skills; 5) not focus on the user as a person with disabilities, but supports the self image of being a person with abilities; 6) remind the solutions that existed before; 7) the use of the products is possible by the information visible/available at all times. They [4][5] further elaborated on this and presented an adaptation of design for all principles in development of assistive technology for people with dementia.

The existing technologies and systems are often expensive and unsuitable. Choosing the appropriate assistive technology is not always easy and there is a wide range of different technologies that can be adapted and used for people with dementia to cope with the practical problems encountered in daily life [6]. The UCD model advocates a design process that involves users in the whole design process in order to match the product to the user requirements and to increase its practical use [13][7]. This study adopted the UCD approach to investigate the processes, devices, and methods employed in dementia care, especially for lost seeking.

This paper is divided into six main sections. Section 1 provides some background information about the beginning of the project. Section 2 outlines the overview solutions for getting lost or wandering. Section 3 describes the interview with questionnaire. Section 4 summarizes results and analysis. Section 5 delineates some of the problems encountered. Finally, the implications of findings are discussed.

II. OVERVIEW SOLUTIONS FOR GETTING LOST OR WANDERING

A. Wandering Behavior

The main character of dementia is the decline of their memory and learning comprehension [1]. Due to the problem of memory and orientation, wandering behaviors and easily-to-be-lost always happens on dementia-sufferers [18]. There are 37% of dementia-patients developing to have wandering behaviors [3]. When they are lost, it is very important to help them and prevent them from wandering. Today, to the family members of the missing elders, they can only report to the Seeking Center of Missing Elders, police, newspapers, mass media, and broadcast or post the missing person's photos.

 ISBN: 978-1-61208-167-0

Table I gives an overview of the solutions for getting lost or wandering behavior as described in the literature. This table is probably not exhaustive. However, the most important solutions are mentioned.

TABLE I. OVERVIEW OF SOLUTIONS FOR GETTING LOST OR WANDERING IN DEMENTIA

Type	Method
Police network	Report to the local police office when family is missing, providing them detailed and complete information, such as their clothing, hair style, blood type, age, gender, height, and obvious characters.
Seeking Center of Missing Elders	Report to the Seeking Center of Missing Elders, connecting to the police network.
Poster	Post posters in public places through the seeking center.
Broadcast media	Ask for audience's help through the Police Radio Station.
Cable station	Announce the seeking information through local cable stations.
Hospital	Inquire the emergency center of local hospitals, or provide them photos to identify.

B. *Lost Seeking Devices*

The consequences of getting lost or wandering are very diverse. Wandering and disorientation may cause them with anxiety, boring, or less exercise [11][16][17]. However, outdoor activity is important to the dementia-patients. Since the patients are level and situation-dependent, the demand from caregivers and individual care-giving environment are different from one by one. Some lost seeking devices with e-Health system are developed and marketed. Most of them are small, easy to wear and carry around, not easily loose, and water-proofing, strong features, but the sizes and usage are different. There are roughly three kind of lost seeking devices: non-electronic, electronic, and biometric (Table II). Each product has its own pros and cons, and can be featured: 1) By the usage: portable, wearable, and biometric; 2) By the materials: paper, plastics, and stainless steel, et al; 3) By the labeled information: photo, name, emergency contact number, address, sufferer's history of chronic cases, battery and system operating signals, and function bottoms.

TABLE II. OVERVIEW OF SOLUTIONS FOR THE LOST SEEKING DEVICES IN DEMENTIA

category	Non-electronic			electronic		biometric
	Self-made ID card	***Case history wrist***	***Missing-preventing bracelets***	***GPS***	***Electronic care system***	***Fingerprint Verification***
						
Usage	Write down personal contact information on a paper card, and carry it around.	Regularly wear on their wrists.	Regularly wear on their wrists.	Wear the signal emitter.	Regularly wear the emitters on their wrists.	Archive fingerprints into the database of the computers in the police office.
Materials	Paper, or covered with films.	Rubber, plastics.	stainless steel	Plastic shell, LED lights.	Plastic shell, Velcro.	None
Labeled information	Photo, name, emergency contact, address.	Name, phone number, Notes of sufferer's history of chronic cases.	Telephone of the seeking center.	LED lights showing battery and operating situation.	Pushing bottoms.	None
Advantages	Portable, simple, and easily made.	Portable, wide variety.	Personal information is registered in the computer of the seeking center. Durable materials.	Care-givers can be aware of the location of the sufferers through computer.	The sufferers can call the care-givers by pushing the bottom.	Employ biometric, no need to wear anything, no worry of losing and labeled.
Disadvantages	Nondurable material; no regular format, cannot be easily found.	The users may have the negative feelings of being labeled and marked.	Unsightly modeling; buckle part is easily loose; The users may have the negative feelings of being labeled and marked.	Expensive; Care-givers must be equipped with computer skills; they need to be careful of the battery condition; easily missing.	Expensive; they need to be careful of the battery condition; use it indoor-only.	Need the identify system and equipment; personal information may leak out, increasing the difficulty of promotion..

 ISBN: 978-1-61208-167-0

III. RESEARCH METHODS

A. Interview with Questionnaire

A status-quo analysis and an interview with questionnaire were conducted. The contents of the questionnaire include subject's personal information, disorientation experience, and their experiences and advices of using the lost seeking-devices. The collected data are analyzed by descriptive statistics. The purpose of the questionnaires is to understand the experience and requirements towards currently available products for dementia-patients. Researchers visited the subjects with the company of social workers. Each interview spent about 40 minutes.

B. Subjects

This study cooperated with the Seeking Center of Missing Elders. Social workers were asked to evaluate the 37 voluntary subjects, who are caregivers and family members of dementia-patients. The qualified subjects are: 1) taking care of adult dementia-patients who are above mild level, 2) taking care of patients having Disability Certificates, or 3) taking care of patients who have experienced disorientation. 17 care-givers in this study are male, 20 female, and their ages were from 28 to 86 years old (M = 50.08, SD = 15.47). The dementia-patients: 16 are male, 21 female, and their ages were from 40 to 93 years old (M = 72.75, SD = 10.23). The education levels of the caregivers are: below junior high school (24.3%), high school (40.5%), college (27%), above graduate school (8.1%); the profession of the caregivers: homemaker (32.4%), business (32.4%), industry (16.2%), free-lance (13.5%), and government employees (5.4%). The relationship between care-givers and patients, the most is the older generation (70.2%), and then spouse (18.9%). The lost seeking devices used by the dementia-patients: the most is self-made ID card (62.1%); the least is case history wrist, GPS and electronic care system. The lost seeking devices which dementia-patients have used: the most is self-made ID card (72.9%); the least is case history wrist, and electronic care system. 16.2% of patients have no experience. The duration of using the lost seeking devices: the longest is 84 months, the shortest is 3 months (M = 28.56, SD = 25.44).

IV. RESULTS AND ANALYSIS

The data is listed in Table III. In order to eliminate the concern of deviation, the responses of this questionnaires are based on care-givers' first-time experience; question 1, 12, 13, 22 are multiple-choice questions; the percentage is calculated on the basis of the number of responses.

TABLE III. RESULTS OF THE INTERVIEW

No	Question	Results and statistics
1	The reason of disorientation?	Walk outside themselves: 26 (46.42 %) Wandering: 9 (16.07 %) Wuditory and visual hallucinations: 8 (14.28 %) Falling down on the road: 4 (7.14 %) Melancholia: 4 (7.14 %) Run away from home: 2 (3.57 %) Suicidal tendency: 1 (1.79 %) Traffic accident: 1 (1.79 %) Mental retardation: 1 (1.79 %)
2	Duration of disorientation?	Less than one hour: 4 (10.8 %) Less than three hours: 8 (21.6 %) Less than five hours: 0 (0 %) Less than twelve hours: 8 (21.6 %) Less than one day: 15 (40.5 %) More than two days: 1 (2.7 %) More than five days: 1 (2.7 %) More than one week: 0 (0 %)
3	How long did you wait until asking for a seeking assistance?	Less than one hour: 13 (35.1 %) Less than three hours: 4 (10.8 %) Less than five hours: 5 (13.5 %) Less than twelve hours: 6 (16.2 %) Less than one day: 3 (8.1 %) More than two days: 1 (2.7 %) More than five days: 0 (0 %) More than one week: 0 (0 %) Never seperated: 5 (13.5 %)
4	Who found your family?	Stranger: 11 (24.3 %) Police: 12 (32.4 %) Family: 5 (13.5 %) Store staff: 3 (8.1 %) Neighbors and chief of village: 3 (8.1 %) Go home by themselves: 3 (8.1 %)
5	Where were they found?	Neighborhood: 14 (37.8 %) Temple: 1 (2.7 %) Hospital: 2 (5.4 %) Park: 8 (21.6 %) Recreation center: 1 (2.7 %) On the road: 4 (10.8 %) Market: 1 (2.7 %) Train station: 3 (8.1 %) Shopping mall: 1 (2.7 %) Go home by themselves: 2 (5.4 %)
6	The distance of disorientation?	Neighborhood: 10 (27%) About 2 kilometers: 14 (37.8%) Neighboring city: 9 (24.3%) Farther than two cities : 4 (10.8%)
7	How was their clothing when there were found?	Intact: 23 (62.1%) Messy and broken: 10 (27%) Nude: 0 (0%) Losing their things: 4 (10.8%)
8	If they were injured when they were found?	None: 31 (83.7%) Abrasion: 5 (13.5%) Scald: 0 (0%) Fracture: 1 (2.7%)
9	Their mental condition when they were found?	Normal: 13 (35.1%) Silence: 18 (48.6%) Mumbling to themselves: 5 (13.5%) Lethargic sleep: 1 (2.7%)

 ISBN: 978-1-61208-167-0

No	Question	Results and statistics
10	Their language expression when they were found?	Normal: 17 (45.9%) Dull: 17 (45.9%) Excited: 3 (8.1%)
11	Their excretory function when they were found?	Normal: 32 (86.4%) Incontinence: 4 (10.8%) Incontinence of feces: 1 (2.71%)
12	Who did you ask for help when the first time your family missing?	Caregiver friend: 2 (4 %) Sufferer's friend: 0 (0 %) Neighbors: 4 (8 %) Chief of village: 1 (2 %) Seeking center: 0 (0 %) Police office: 14 (28 %) Seek on your own: 7 (14 %) Spouse of the care-giver: 7 (14 %) Spouse of the sufferer: 4 (8 %) Children of the sufferer: 5 (10 %) Brotherhood of the care-giver: 6 (12 %)
13	The choice of lost seeking report when the first time missing happened.?	Seek on your own: 35 (64.8%) Post photos and posters: 0 (0%) Instant broadcast by chief of the village: 3 (5.5%) Seeking center: 0 (0%) Police office: 14 (25.9%) Hospital: 1 (1.8%) The city funeral parlor: 0 (0%) Local broadcasting radio: 1 (1.8%) Local television station: 0 (0%)
14	When your family were miss again, is the location they being found the same as before?	Yes: 3 (8.1%) No: 34 (91.8%)
15	Was the way you seek them the same as the first time?	Yes: 34 (91.8%) No: 3 (8.1%)
16	Did you buy the seeking-assist products?	Yes: 21 (56.7%) No: 16 (43.2%)
17	What kind of seeking-assist products you have used before?	No experience: 6 (16.2%) Self-made ID card: 27 (72.9%) Case history wrist: 0 (0%) Bracelets: 2 (5.4%) GPS: 2 (5.4%) Electronic care system: 0 (0%)
18	What kind of seeking-assist products you are using currently?	Self-made ID card: 23 (62.1%) Case history wrist: 0 (0%) Bracelets: 14 (37.8%) GPS : 0 (0%) Electronic care system: 0 (0%)
19	The reason?	Doctor's suggestion: 6 (10%) Having disoriented experience: 11 (18.3%) Chronic disease: 6 (10%) Wandering behavior: 13 (21.6%) Suicidal tendency: 0 (0%) Suffering from melancholia: 1 (1.6%) Auditory hallucinations or illusions: 1 (1.6%) Language disorders: 2 (3.3%) Mental retardation: 2 (3.3%) They wore it spontaneously: 12 (20%) Police's suggestion: 6 (10%)
20	How many times of missing before you bought the products?	1: 15 (40.5%) 2: 7 (18.9%) 3 ~ 5: 8 (21.6%) More than 5: 7 (18.9%) (M = 3.5, SD = 4)
21	The acceptance of your family to the product?	High acceptance: 3 (8.1%) Acceptable: 14 (37.8%) No comment: 6 (16.2%) Kind of reject: 5 (13.5%) Totally reject: 8 (21.6%)
22	Who do you ask for help when your family missing with using the products?	Care-giver friend: 2 (4 %) Sufferer's friend: 0 (0 %) Neighbors: 4 (8 %) Chief of village: 1 (2 %) Seeking center: 17 (25.3 %) Police office: 14 (28 %) Seek on your own: 7 (14 %) Spouse of the care-giver: 7 (14 %) Spouse of the sufferer: 4 (8 %) Children of the sufferer: 5 (10 %) Brotherhood of the care-giver: 6 (12 %)
23	The seeking method you chose when your family missing again with using the products?	Seek on your own: 12 (24.3%) Post photos and posters: 0 (0%) Instant broadcast by chief of the village: 3 (8.1%) Seeking center: 4 (10.8%) Police office: 21 (56.7%) Hospital: 0 (1.8%) The city funeral parlor: 0 (0%) Local broadcasting radio: 0 (1.8%) Local television station: 0 (0%)
24	Which way of seeking you think would be the most efficient?	Help myself: 5 (13.5%) With the assistance of police: 6 (16.2%) Stranger's concern: 7 (18.9%) ID documents: 6 (16.2%) With the assistance of technology: 13 (35.1%)

 ISBN: 978-1-61208-167-0

No	Question	Results and statistics
25	Which kind of seeking-assist products you will use in the future?	Self-made ID card: 5 (13.5%) Case history wrist: 1 (2.7%) Bracelets: 25 (67.5%) GPS: 3 (8.1%) Electronic care system: 3 (8.1%)
26	Do you agree with applying high-tech to seeking-assist facility for demented elders?	Yes: 33 (89.1%) No: 4 (10.8%)
27	Are there any concern of leakage of personal information on the application of high-tech to seeking-assist facility for demented elders?	Yes: 15 (40.5%) No: 22 (59.4%)

Some advices obtained from subjects in the interviews can be summarized in Table IV.

TABLE IV. SOME ADVICES OBTAINED FROM SUBJECTS IN THE INTERVIEW

Advice	No. of Subjects
Dementia-patients move slowly, but would disappear if the care-givers and family members didn't keep their eyes on the sufferers.	17
Prepare some kind of ID cards for the dementia-patients. Some of them are using more than two kinds of lost-seeking devices.	13
Though the phone number of the seeking center is shown on the product, it is not clear and hardly be trusted.	5
Anything related to computer is too complicated to ;learn and is untrusted.	7
Dementia-patients tend to play, fiddle with, or dismantle the devices, which cause the products loosing or missing.	5
Dementia-patients may have the negative feelings of being labeled, so they tend to resist or forget to wear the devices and then lose them.	10
Dementia-patients, when they are lost, tend to lose their things, including the lost-seeking devices.	6
Subcutaneous implantation of microchips and biometric would become a way of lost-seeking.	4
Electronic products are relatively expensive and constrained by the signal reception.	9
The products can only provide limited help.	15

V. CONCLUSION AND DISCUSSIONS

Care-givers' opinions towards lost-seeking report and seeking-assist products were discussed as follows: 1) Reporting to police would be the best choice: when the dementia-patients are missing, care-givers tend to seek on their own, and most of them ask for police's help. With the time passing, police office will announce seeking notice to the seeking center. Thus seeking center is not in the front line of lost-seeking, but working as an information manager. 2) Care-givers' computer skill should be concerned: if the care-givers are elders, or in a lower education level, their computer skills are limited. They showed higher resistance towards electronic and GPS monitoring systems. 3) The seeking-assist product is not colorful enough: probably because too less propaganda or the color of the products is not obvious enough, or they wear the products in a wrong way, when the patients are found on the road, the passengers may be not aware of the patients were wearing seeking-assist products. 4) The patients show resistance to the products: care-givers indicate that the patients showed resistance towards the seeking-assist products, especially male patients. Female patients tend to consider the products as accessories, showing higher acceptance towards bracelets product. Care-givers usually persuade the patients in a way of religion or family love, warning them not to take off the products. 5) The materials of the products are not strong enough: some patients tend to play, fiddle with, or dismantle the products, which cause the products loose or miss. Also, when the products are worn-down or eroded by water, care-givers tend to replace or repair the product by themselves, which causes the specification more messy. 6) Concern of labels and data leakage: most care-givers think the data leakage is not a big deal if the patients can be rapidly found. They indicate that the identification of labeling should be enhanced.

Table III shows there are 46.42% of dementia-patients missing because they walk outside themselves, accompanying with wandering (16.07%), auditory and visual hallucinations (14.28%). Duration of disorientation is mostly less than one day (40.5%), and the care-givers would ask for seeking assistance when the patients are missing within one hour (35.1%). 48.6% of patients were silent when they were found; 45.9% are dull; some old patients were with abrasion due to falling down; incontinence may be caused by tiredness. Therefore, they should be psychologically comfortable treated to relieve their pressure. Regarding the person care-givers ask for help, police is credited with 32.4%, and passengers are 24.3%. Also, the care-givers tend to seek by themselves or by their family, if they are not able to find the patients, they will report to police office and ask for their help, which is in accordance with the response in question 12 and 13. Comparing the disoriented distance, mostly are in the neighborhood and less than 2 kilometers (37.8%). 24.3% patients were found in the neighboring city. However, in the responses of the choice of lost-seeking (question 12 and 13), chief of the village and instant broadcasting are not their first choice. We presume that they might not know these methods, or they take it as privacy, and dislike to be known by others know. In the question of 14 and 15, when the patients get lost again, 91.8% are found in a different location, 91.8% of care-givers seek the patients in the same way as the first time, which shows the route and situation are uncertain, seeking in the same way would reduce the efficiency. Perhaps the

 ISBN: 978-1-61208-167-0

situation of the patients is uncertain, and the care-givers have only limited way of seeking. This is a must-solve problem, and it is also the main issue of this study.

In the question of what products care-givers have used and what they are using, 16.2% are first-using. Most people use non-electronic products, self-made card (61.1%) and bracelets (37.8%) are the top 2 choices, and some care-givers say they would use both at the same time. For open question, they show that there is telephone number of the seeking center in the bracelets; if they cannot find the card on the patients, bracelets would be the last defensive line. Electronic products are with higher price, and computer skills are required. Considering that 24.3% of care-givers are with the education level of below junior high school, 40.5% are high school, 18.9% are spouses, and most of them are elders, it may be the reason that they tend not to use the electronic products. Additionally, Case history wrist is the least used; it may be because the materials have less durability, even though there is contact information on it. In the question 19 and 20, the reason of using the product is wandering behavior (21.6%), which is in accordance with question 1. Furthermore, most care-givers bought the products after the 3.5 times of missing experience, but only 20% family members would buy the products spontaneously. In addition, the suggestion from doctors (10%) and police (10%) would also be the main reason, which may because of their professional image.

As for the acceptance toward the products, 37.8% of patients show high acceptance, but there are also 21.6% showing totally rejection, often taking the products off or rejecting to wear. Since most patients are the older generation, care-givers persuade them by religion and family love, which may be the reason. Comparing question 22 and 23: care-givers' choices of seeking-assistant and seeking method before and after using the products: 25.3% would choose seeking center, 56.7% choose police office, but the number of seeking themselves decreasing (referring to question 12 and 13), which may because the care-givers mainly use bracelets as the seeking product, and police office is the first choice for reporting. We hope the seeking center can become a center for consultation, the products did relieving care-givers' burdens. Referring to question 24, 25, 26, and 27, technology-assist (35.1%) would quickly find out the missing patients. However, bracelet (67.5%) is the first choice of the future product; 89.1% agree with the application of technology-assist product on aged patients, but there are also 59.4% consider that it may cause data leakage of personal information. We infer that most care-givers hope technological products would increase the efficiency and safety, but they also think it's too expensive and lack of computer skills. This result demonstrates the choice of seeking methods depends on the education level of the care-givers and most of them are elders. The concern of data leakage is also related to today's fraud issue, which may be the reason limiting the promotion of electronic products and biometrics.

Based on the interview and analysis mentioned above, some problems are concluded in Table IV: 1) Dementia-patients move slowing, but disappear when the care-givers and family members didn't keep their eyes on the sufferers. 2) Most care-givers and family members prepare cards for the dementia-patients; some of them are using more than two kinds of seeking-assist products. 3) People are not familiar with the seeking-assist products: they don't know the phone number of the seeking center is on the product. 4) Computer skills are needed for some electronic products. It is difficult to those care-givers or family members who are not well-educated or are in a ripe old age. 5) Some dementia-patients tend to play, fiddle with, or dismantle the products, which cause the products loose or miss. 6) Dementia-patients may have the negative feelings of being labeled and marked, so they tend to resist or forget to wear the products and then lose them. Some care-givers and family members entice the patients to wear the products by their regions, warning them not to take down or miss the products. 7) Some dementia-patients, when they are lost, tend to lose their things, including the seeking-assist products. 8) Most of the seeking-assist products are small, and easily loose, but the sizes and usage are different, and there are also 21.6% Dementia-patients showing totally rejection, often taking the products off or rejecting to wear. Some care-givers and family members expect subcutaneous implantation of microchips and biometric would become a way for seeking. Comparing question 26 and 27: 89.1% care-givers and family members agree with applying high-tech to seeking-assist facility for demented elders. But 59.4% care-givers and family members disagree any concern of leakage of personal information on the application of high-tech to seeking-assist facility for demented elders. 9) Since electronic products are relative expensive and limited by its signal reception, the users are very few; those care-givers and family members with higher education level say they hope the ability of GPS would improve, and its price would decrease. 10) Most care-givers and family members say that for moderate and severe dementia-sufferers, and those who can walk around themselves, the products can only provide limited help. They will finally take the patients to a sanatorium, or hire a care-giver to take care of the patients.

The problems concluded above can be categorized into two directions: disorientation experience and lost seeking. The first one is about Problem (1), (3), and (7), which are related to disorientation experience and choosing a way of seeking; the second one is about Problem (2), (4), (6), (8), (9), and (10), which are the common problems of both ways and products for seeking.

VI. SUGGESTION

From the above discussion, we propose that the seeking products and methods should be re-examined and redesigned. More researches should be implemented in order to increase the sufferers' preference to use the products, and reduce the burden of the care-givers. The general principle of seeking method and product design is suggested as follows: 1) Identification of the products: the products should be worn at a clear position and with vivid color, which would help people quickly find the contact information and increase the efficiency. 2) Since the youngest care-giver was 40 years old, and the average age of subjects was 72 years old, their

 ISBN: 978-1-61208-167-0

computer skills should be put in concern, if telemedicine service is implemented. 3) The acceptance of the product: reducing the resistance of dementia-patients, and designing the products not easily to dismantle. We should also consider the habits and types of male, female, and patients in different levels. 4) Development of the local reporting system in the community: Since dementia-patients usually get lost in their neighborhood, and there are also 25.3% care-givers and family members would choose seeking center, 56.7% choose police office, a network connecting the office of village chief, recreation center, police office, the seeking center, and an online toll free assistance centre should be set up for convenience locating the lost. 5) Improving the lost-seeking poster design and promotion: according to the collected data, the sufferers were usually found by passengers, through promotion and lost-seeking poster, people will pay more attention on the missing patients. 6) After the lost were found, some actions and consultants should be adopted in order to comfort the patients and reduce their mental pressure in time.

Not only dementia-patients, this principle can also be applied to disabled elderly, mental disorders, and mental retardation patients. The seeking-assist product is still not common; we should do more promotion and lead more people to use the seeking-assist products. Also, the development of reporting system can decrease the burdens of care-givers and police officers. The seeking center would play the role of consultation, report, and promotion. We also expect people will pay more attention to the dementia-patients on the road, guaranteeing their safety.

VII. LIMITATIONS

This paper discovers the problems of current seeking method and facilities, investigating 37 care-givers' opinions, and concluding some suggestions. However, it is very difficult to have subjects willing to take the test. The subjects are mostly from the Seeking Center of Missing Elders, and they mostly live in Taipei City and Taipei Country. It means they may not be the representative of the target population. Therefore, the conclusion of this study probably can only be applied to the Taipei metropolitan. We will consider other towns into our research and further discuss the opinions of care-givers towards seeking methods and facilities.

ACKNOWLEDGEMENT

Thanks for the cooperation of the Seeking Center of Missing Elders, and the financial support from the ROC National Science Counsel (NSC99-2221-E-036-032), and the Tatung University (B99-D01-071).

REFERENCES

[1] D. L. Algase, "Wandering in dementia," Annu Rev Nurs Res, vol. 17, 1999, pp. 185–217.

[2] K. D. Bail, "Electronic tagging of people with dementia: devices may be preferable to locked doors," BMJ, vol. 326, no. 7383, 2003, pp. 218.

[3] C. G. Ballard, R. N. C. Mohan, and C. Bannister, "Wandering in dementia suffers," Int J Geriatr Psychiatry, vol. 6, 1991, pp. 611–614.

[4] S. Bjørneby, P. Duff, and O. M˜aki, "Developing assistive technology for people with dementia," Assistive Technology – Shaping the Future, G Craddock, ed., IOS Press, 2003, pp. 781–786.

[5] S. Bjørneby, P. Topo, and T. Holthe, "Technology, ethics and dementia," Norwegian Centre for Dementia Research, 1999.

[6] S. Brianna, B. A. Fjeldsoe, L. M. Alison, and D. M. Yvette, "American Journal of Preventive Medicine," vol. 36, no. 2, 2009, pp. 165-173.

[7] R. D. Buurman, "User-centered design of smart products. Ergonomics," vol. 40, 1997, pp. 1159−1169.

[8] S. Cahill, "Technologies may be enabling," BMJ, 2003, pp. 281.

[9] G. Demiris, L. B. Afrin, S. Speedie, K. L. Courtney, M. Sondhi, V. Vimarlund, C. Lovis, W. Goossen, and C. Lynch, "Patient-centred applications: use of information technology to promote disease management and wellness," Journal of the American Medical Informatics Association, vol. 15, no. 8, 2008, pp. 121-126.

[10] C. P. Ferri, M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, M. Ganguli, K. Hall, K. Hasegawa, H. Hendrie, Y. Huang, A. Jorm, C. Mathers, P. R. Menezes, E. Rimmer, and M. Scazufca, "Global prevalence of dementia: a Delphi consensus report," Lancet, vol. 366, 2005, pp. 2112–2117.

[11] B. Fjeldsoe, A. Marshall, and Y. Miller, "Behavior change interventions delivered by mobile telephone short-message service," American Journal of Preventive Medicine, vol. 36, no. 3, 2009, pp. 165-73.

[12] J. C. Hughes, and S. L. Louwe, "Electronic tagging of people with dementia who wander," BMJ, vol. 325, 2002, pp. 847–884.

[13] T. Kontogiannis, and D. A. Embrey, "User-centred design approach for introducing computer-based process information systems," Appl Ergonomics, vol. 28, no. 2, 1998, pp. 109 – 119.

[14] I. Leroi, and C. G. Lyketsos, "Neuropsychiatric aspects of dementia. in: Dementia," A. Burns, J. O'Brien, D. Ames, and H. Arnold, eds, 2005.

[15] M. Mapstone, T. M. Steffenella, and C. J. Duffy, "A visuospatial variant of mild cognitive impairment: getting lost between aging and AD," Neurology, vol. 60, 2003, pp. 802–808.

[16] R. Passini, C. Rainville, N. Marchand, and Y. Joanette, "Wayfinding in dementia of the Alzheimer type: planning abilities," J Clin Exp Neuropsychol, vol 17, 1995, pp. 820–832.

[17] R. Passini, C. Rainville, N. Marchand, and Y. Joanette, "Wayfinding in a nursing home for advanced dementia of the Alzheimer's type," Environment and Behavior, 2000, pp. 684-707.

[18] S. M. C. Rasquin, C. Willems, S. de Vlieger, R. P. J. Geers, and M. Soede, "The use of technical devices to support outdoor mobility of dementia patients," Technology and Disability, vol. 19, 2007, pp. 113-120.

 ISBN: 978-1-61208-167-0

Service Control System Based on Behavioral Characteristics of User Multitasking

Satoshi Uemura, Sumaru Niida, Hajime Nakamura
Communications Network Planning Laboratory
KDDI R&D Laboratories Inc.
2-1-15 Ohara Fujimino-shi Saitama 356-8502 JAPAN
{uemura, niida, nakamura}@kddilabs.jp

***Abstract*—With the rapidly growing deployment of high performance mobile terminals, such as smart phones, multitasking becomes one of the popular uses of mobile services. As the user behaviors become more diverse, a service control based on user's behavior would be required. In general, users do not always pay attention to all services in terms of the progress of those services while multitasking. Thus, service control can be expected to maintain user satisfaction with the service at high levels by allocating computer and network resources to the specific service to which the user's attention is directed. This paper proposes a service control system based on the behavioral characteristics of users while multitasking. As an initial study for establishing a service control system, we first investigated the behavioral characteristics of users while multitasking through a subjective assessment of 400 participants. Furthermore, we verified the proposed system for the effect on traffic reduction by numerical simulations with the behavioral data and confirmed the effectiveness of the proposed system. The main contribution of this paper is that we clarified that the service control system based on behavioral characteristics would be useful in constructing a cost-effective network.**

Keywords-multitasking; user satisfaction; mobile service; service control

I. INTRODUCTION

Thanks to the rapid growth of mobile Internet services, we enjoy a variety of services everywhere whenever we want. Mobile services would be expected to expand continuously by diversifying according to user demand and behavior [1]. For telecommunication carriers and service providers, providing mobile services that ensure user satisfaction is one of the most crucial issues. User satisfaction with a service is called Quality of Experience (QoE). The term QoE is defined in ITU-T as follows [2]:

Quality of Experience (QoE): The overall acceptability of an application or service, as perceived subjectively by the end-user.

Notes

1) Quality of experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.).

2) Overall acceptability may be influenced by user expectations and context.

The QoE assessment for the mobile communication service has been developed in the audiovisual research field, such as telephone speech and video. ITU-T P.800 [3] and ITU-R BT.500 [4] are representative. Recently, the QoE assessment for such interactive services as Web browsing and e-mail has been studied. Although many factors, such as reliability of the service and ease of use, can be expected to affect the QoE for interactive services, the most dominant factor was reportedly speed [5]. Traditionally, the so-called 8-second rule [6] was used as the index for Web site planning. ITU-T G.1030 [7] or other assessments [8] are currently recommended to determine the appropriate index. Based on the idea that subjective assessments with a method having high-ecological validity are important, the authors proposed an assessment method to assess the QoE under the actual context of use by using Web scripts [9].

Meanwhile, with the rapidly growing deployment of high performance mobile terminals, such as smartphones, multitasking becomes one of the popular uses of mobile services. Although the amount of traffic can be expected to continuously increase, in the mobile communication environment, the computational resources of mobile terminals and network resources are strictly restricted. Increasing network capacity to cope with huge volumes of traffic is not desirable because of the cost. Therefore, the available resources should be used effectively to maintain user satisfaction with services at high-level by implementing some sort of resource allocation and service control. In order to construct such a cost-effective network, we propose a service control system based on the behavioral characteristics of users while multitasking. In general, users do not always pay close attention to all services in terms of the progress of those services during multitasking. Thus, the behavioral data of users while multitasking would be useful information in the implementation of some sort of service control. We confirmed the effectiveness of the proposed system by numerical simulation using behavioral data obtained by subjective assessments with 400 participants.

This paper is organized as follows. In Section II, resource allocation methods to improve service quality are summarized. In Section III, we propose a service control system based on behavioral characteristics of users while multitasking. In Section IV, the field evaluation results regarding user satisfaction with services and the behavioral characteristics while multitasking are given. We also verify the effectiveness of the proposed system through a numerical

 ISBN: 978-1-61208-167-0

simulation of bandwidth use based on the behavioral data obtained by the subjective assessment.

II. RELATED WORK

Here we briefly explain the conventional methods to improve service quality. As described previously, allocating limited resources is one of the most vital factors in the mobile environment. The basic approach is to allocate resources in order to maximize some sort of function.

Utility-based resource allocation methods have been widely studied [10]-[12]. The purpose of the utility-based approach is to maximize total utility defined by the utility function. Thus if the utility function is used that consists of metrics that reflect the user's perception for a certain service, then the function can be expected to maximize the QoE for the service. However, since the user's perception of a certain service can be affected by his/her circumstances and the context of use [13], it is difficult to allocate the resources, so as to maximize the QoE.

Context-aware resource allocation has received much attention recently [14]-[16]. Proebster et al. proposed a context-aware resource allocation method to improve the QoS of heterogeneous traffic [16]. By using context information collected from user terminals, the base station scheduler can allocate the available resources to each user more efficiently than previous methods. For example, this method focuses on the foreground/background state of the service/application generating the traffic and allocates more resources to the service in the foreground. The foreground/background state for a certain service was previously determined using this method, e.g., web browsing is the foreground and file downloading is the background. However, in a multitasking environment, a certain service/application can be used as the foreground sometimes and as the background at other times. The foreground/background state and the context of use for a service/application can vary from time to time and the timing of the change is different for each user. Therefore, we need resource allocation or a service control method that can handle the variations of the context of use for a certain service immediately.

III. SERVICE CONTROL SYSTEM BASED ON BEHAVIORAL CHARACTERISTICS OF USERS MULTITASKING

We propose a service control system based on behavioral characteristics of users while multitasking. In a multitasking environment, although users can run a variety of services and applications simultaneously, they do not always focus their attention on every service and application. User satisfaction with a service fluctuates depending on the style and context of use. In a multitasking environment, the service can be enjoyed as either the foreground (FG) task or the background (BG) task. When a certain service is enjoyed as a BG task, the focus of the user's attention would be relatively low compared with when the service is enjoyed as a FG task. Therefore, the user can be expected to accept the given service quality for the BG task even when the quality is not very high. If we can control the network and server resources appropriately based on the behavioral characteristics as long as the information about the application type and the context of use collected from each user, the service control method can be expected to retain user satisfaction as a high level. For example, suppose a large number of users enjoy a certain service. Their styles of use would be different; some users use the service as a FG task and others use the service as BG tasks. In this case, if the service quality for the user of the service as a BG task was degraded by implementing some sort of service control, the user satisfaction for the service may not deteriorate.

In order to establish such a service control method based on the behavioral characteristics of users, one possible system configuration is shown in Figure 1. Each user terminal monitors the context of use regarding the application running on the terminal and behavioral data of the user and sends the monitoring results to the service control server (SCS) at the appropriate time. The SCS determines the service control policy based on the collected information regarding each user terminal and sends the determined service control policy to the user terminals, the intermediate nodes, and the Web servers depending on the policy. The user terminal can control the computational resource allocation with each task. The intermediate node and the Web server can implement priority control and traffic shaping, e.g., changing the bandwidth allocation.

In addition to the above service control method that implements traffic control on the user terminals and intermediate nodes, it is possible to implement the QoE improvement method that provides an appropriate time filler for the user [17]. The typical time fillers are a loading bar, a progress bar, and other short movies. These prearranged fillers can be stored on the terminals by default, and the terminal then displays them during the waiting process. Additionally, more flexible news, trivia, and advertisements can also be utilized as time fillers. In [17], we confirmed that the time filler has a significant effect on the decrease in dissatisfaction while waiting, though it is strongly influenced by context.

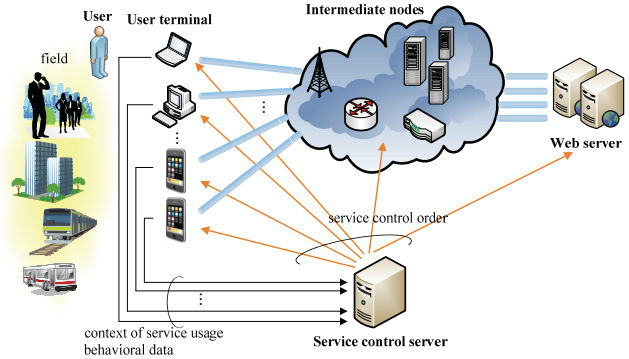


Figure 1. System configuration of the proposed method.

IV. EXPERIMENTS

In order to establish the proposed service control system, we first have to clarify the behavioral characteristics during multitasking and user satisfaction with the service in both cases where the service is running as a FG task and a BG task. Here we conducted the subjective assessment of mobile services by using the high ecological field evaluation method proposed in [9].

A. Field evaluation system

The overview of the proposed Web script-based method is depicted in Figure 2. As shown in Figure 2, our proposed method is extremely simple, since only a subjective assessment server (SAS) is required to construct the system. The SAS plays two roles: distributing a program for artificial context (PAC) written in a Web script such as FlashLite and receiving the answers from participants. In our proposed method, the participant's own mobile terminal can be used without any modification, since most commercial mobile terminals support Web scripts by default. Thus, a large-scale subjective assessment can be conducted easily in the field with little effort, since the test organizer needs to prepare neither customized terminals, nor elaborate network systems.

A general sequence flow for conducting the subjective assessment using this method is as follows. The participant first sends a request to the SAS via the Internet. Then, as a response, the SAS distributes the PAC to the mobile terminal. Second, the participant conducts the field evaluation using the PAC. As described later, since the PAC can be run without an actual communication environment, the participant can conduct the field evaluation anywhere. Finally, the participant sends the results to the SAS via the Internet. Once the participants have obtained the PAC, all necessary procedures can be accomplished according to the instructions for the PAC.

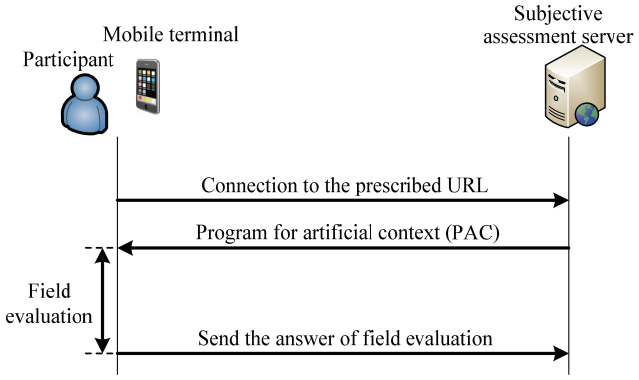


Figure 2. Sequence flow for conducting the subjective assessment.

B. Behavioral characteristics on multitasking

We conducted a subjective assessment to evaluate user satisfaction with the waiting time while content is downloading using smartphones. In this test, participants used their own smartphones and assessed the waiting time during downloading with a five-grade rating score. The PAC utilized in this test is an application for an Android terminal and shown in Figure 3. The PAC emulates the tabbed browser. Here we assume the situation where the user is Web browsing in tab A while downloading a video file in tab B. Although tab A is the foreground task and tab B is the background task at first, the participants can switch the FG/BG task freely by touching the tab. Note that the download speed is not changed regardless of whether the tab B is a foreground or background task. The time when the participant switched the FG/BG task is logged with the PAC as data representing user behavior.

By touching the start button of Picture #1 in Figure 3, the displayed picture immediately changes into Picture #2, which shows the Web browsing in tab A and the downloading in tab B. After a predetermined waiting time, Picture #2 changes into Picture #3 and reporting that downloading has finished; note that even when the downloading has finished, tab B does not automatically become the foreground. Thus, the participant makes tab B the foreground task by touching tab B and goes ahead with the assessment by touching the Next button in Picture #4. In Picture #5 of Figure 3, the participant evaluates the waiting time with a five-grade rating score.

The experimental conditions are shown in Table I. As shown in Table I, in tab A, trivia consisting of about 150 words or a *senryu* consisting of 17 words are shown; note that the *senryu* is a traditional form of Japanese poetry with a 5-7-5 syllable structure. Although there are individual differences, it needs about 12 seconds and 2 seconds to read through the trivia and the *senryu*, respectively. In conditions 1 and 3, the progress bar that represents the progress of the download is shown in tab B. The user interface design for each condition is shown Figure 4. In each condition, the participant assesses 5 patterns of waiting time. We recruited a total of 400 participants consisting of 224 males and 176 females; note that all participants conducted the subjective assessment using their own smartphones.

The subjective assessment results are shown in Figure 5. As seen in Figure 5, user satisfaction with the waiting time is different depending on the conditions. We tested the differences between the conditions by applying two-way ANOVA with two independent variables: waiting time and conditions. Accordingly, it was found that the main effect of condition was significant at the 1% level. By comparing conditions 1 and 2 and conditions 3 and 4, we can see that by representing the progress bar, user satisfaction with the waiting time improved. For both conditions 1 and 2, user satisfaction with waiting time improved compared with condition 5, while in conditions 3 and 4, user satisfaction deteriorated. In conditions 1 and 2, user attention was focused on reading the trivia, and they were not aware they were waiting for the download to finish. Consequently, user satisfaction did not deteriorate.

Next, the number of times the FG/BG task switched during one given condition is shown in Figure 6; the number of switches was calculated as the average of all participants. As it can be seen from Figure 6, the participants tended to switch FG/BG task more frequently when the progress bar was not shown in tab B. We performed the two-way ANOVA with two independent variables, waiting time and

 ISBN: 978-1-61208-167-0

condition, then it was found that both the main effects of waiting time and condition were significant at the 1% level.

Next, we classified the participants into two groups per each condition. The participants who switched FG/BG task at least one time during each condition were classified into the first group. The other participants were classified into the second group. Figure 7 shows the user satisfaction of each group. Note that in conditions 1 and 3, since the number of participants who switched FG/BG task at least one time during the given condition was few, we did not analyze per each group any more. As can be seen from Figure 7, the participants who frequently switched FG/BG task tended to be dissatisfied with the waiting time.

Consequently, the users who used the service as a BG task were not dissatisfied with the service even though the service quality was degraded. Therefore, it can be expected that the service control system would work well if we used the behavioral data of the user appropriately.

TABLE I. Experimental conditions

Condition	Tab A	Tab B	Waiting time [sec]
1	Trivia	Progress bar	4, 10, 15, 20, 25
2	Trivia	--	4, 10, 15, 20, 25
3	Senryu	Progress bar	4, 10, 15, 20, 25
4	Senryu	--	4, 10, 15, 20, 25
5	--	--	4, 10, 15, 20, 25

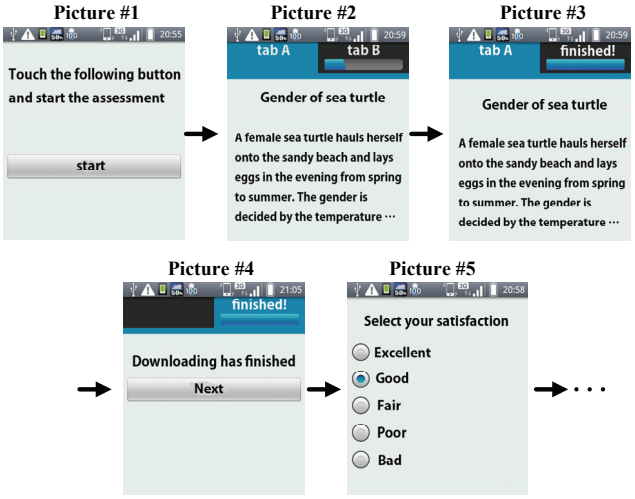
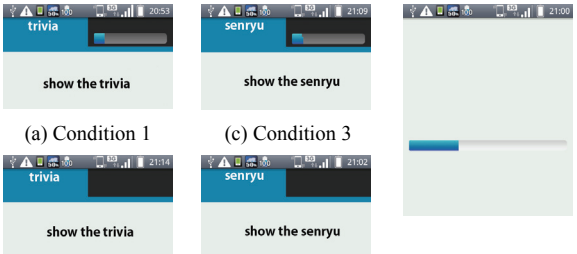


Figure 3. Example of screen transition in the PAC.



(a) Condition 1 (c) Condition 3

(b) Condition 2 (d) Condition 4 (e) Condition 5

Figure 4. User interface design of each condition.

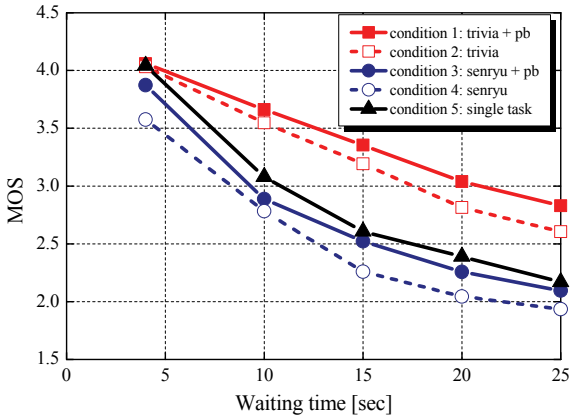


Figure 5. The subjective assessment results.

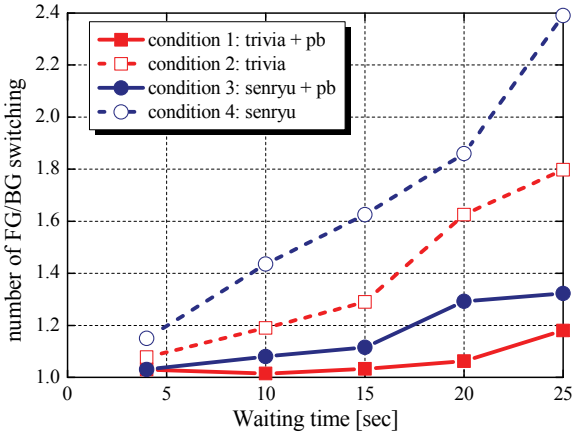


Figure 6. The average number of FG/BG switches per condition.

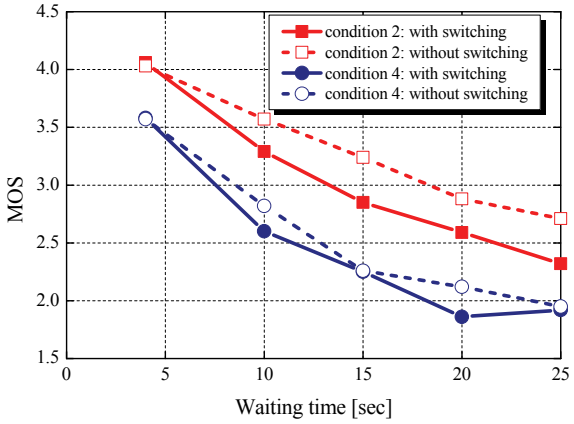


Figure 7. The relationship between user satisfaction and number of switches.

 ISBN: 978-1-61208-167-0

C. *Numerical simulation*

Here we verify the effectiveness of the proposed service control system by a numerical simulation based on the behavioral data obtained from the subjective assessment in the previous subsection. We simulated how much traffic could be reduced by using the proposed system. In this simulation, we assumed that about a thousand users browsed the Web while downloading a certain content file. The experimental conditions are shown in Table II. As shown in Table II, we set the content file size to 1.0 MB. As seen in Figure 5, if we adopt the waiting time when the reference value of MOS is set to 3.0, the acceptable waiting time for condition 1 is about 20 seconds and that for condition 5 is about 10 seconds. Therefore, we assumed that the acceptable waiting time when the user uses the service as a FG task was 10 seconds and that for the BG task was 20 seconds. This meant that the required throughput was 800 kbps when the content file download was performed as a FG task, while the required throughput was 400 kbps when the download is a BG task.

As described in the previous subsection, we obtained the behavioral data regarding the FG/BG switching timing of 400 participants. By using the behavioral data of condition 1 in the subjective assessment, we generated a traffic generation pattern for each user as shown in Figure 8. For example, since user #1 downloaded the file as a BG task to the last, it followed that user #1 required throughput of 400 kbps to the last. Since some users, such as the users #2 and #3, downloaded the content as a FG task occasionally, it followed that they required throughput of 800 kbps occasionally. Note that we did not care that there were identical traffic generation pattern among the 400 patterns.

Here we use the 3 different access patterns shown in Figure 9; access patterns (a), (b), and (c) follow a normal distribution in which the mean and standard deviation are 20 and 7 seconds, 20 and 5 seconds, and 20 and 2 seconds, respectively. The number of total users in each access pattern is 974, 979, and 991, respectively. In each access pattern, we randomly selected one participant among 400 participants and generated the traffic pattern for the selected participant shown in Figure 8.

Figure 10 shows the simulation results regarding the required bandwidth. As can be seen from Figure 10, the required bandwidth can be reduced by using the proposed system. The reduction rate for access pattern (c) is larger than that for access pattern (a). Thus, the proposed service control system worked well because access timing was more concentrated in short period of time. User satisfaction with the service would improve if we allocated redundant bandwidth to the user who needed more bandwidth.

TABLE II Experimental conditions

Parameter	Value
Content file size	1.0 MB
Required throughput	400 kbps (for BG task) 800 kbps (for FG task)

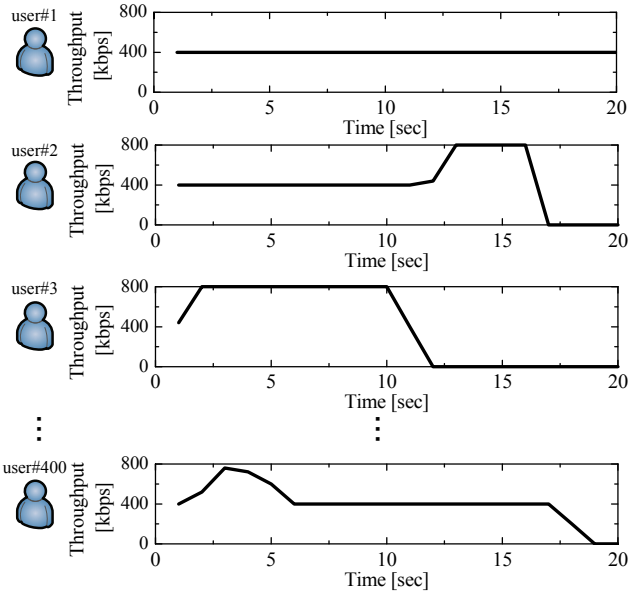


Figure 8. Example of traffic generation pattern.

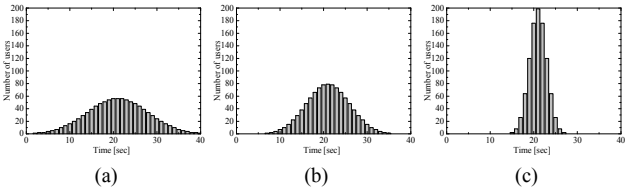


Figure 9. Access pattern utilized in the experiment.

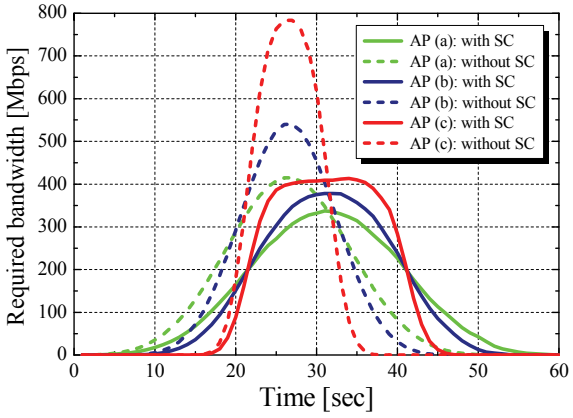


Figure 10. Simulation results of required bandwidth.

V. CONCLUSION

This paper presented a service control system based on behavioral characteristics of users while multitasking. As an initial study toward establishing the service control system, we first assessed user satisfaction with waiting time when downloading and the behavioral characteristics while multitasking by conducting a subjective assessment with 400

 ISBN: 978-1-61208-167-0

participants. As a result, users tended to accept relatively long waiting times when they used a certain service as a background task.

Next, we verified the effectiveness of the proposed service control system using a numerical simulation based on behavioral data obtained from the subjective assessment. Consequently, the more the access timing was concentrated within a short period of time, the greater the effect on traffic reduction.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Prof. E. Harada of Tsukuba University for her support of our work.

REFERENCES

[1] ITU-R Report M.2072, "World mobile telecommunication market forecast," International Telecommunication Union, Nov. 2006.

[2] ITU-T Recommendation P.10 Amendment 2 "Vocabulary for performance and quality of service," International Telecommunication Union, Jul. 2008.

[3] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality," International Telecommunication Union, Aug. 1996.

[4] ITU-R Recommendation BT.500-11, "Methodology for subjective assessment of the quality of television picture," International Telecommunication Union, June 2002.

[5] A. Bouch, M.A. Sasse, and H. DeMeer, "Of packets and people: a user-centered approach to quality of service," proc. IEEE International Workshop on Quality of Service (IWQoS 2000), pp. 189-197, June 2000.

[6] Zona Research, "The Need for Speed II," Zona Market Bulletin, no.5, Apr. 2001.

[7] ITU-T Recommendation G.1030, "Estimating end-to-end performance in IP networks for data applications," 2005.

[8] E. Ibarrola, F. Liberal, I. Taboada, and R. Ortega, "Web QoE Evaluation in Multi-Agent Networks: Validation of ITU-T G.1030," Proc International Conf. on Autonomic and Autonomous Systems (ICAS 2009), Valencia, April 2009.

[9] S. Uemura, S. Niida, and H. Nakamura, "A Web script-based method to evaluate user satisfaction of mobile services in the field," Proc. CENTRIC 2010, pp. 34-39, Aug. 2010.

[10] T. Kaji, Y. Minoda, K. Tsukamoto, and S. Komaki, "Radio agent realizing user utility-based resource allocation in wireless spot access," International symposium on communications and information technologies (ISCIT), Oct. 2004.

[11] W. Kuo and W. Liao, "Utility-based radio resource allocation for QoS traffic in wireless networks," IEEE Transactions on wireless communications, vol.7, No.7, Jul. 2008.

[12] M. Katoozian, K. Navaie, and H. Yanikomeroglu, "Utility-based adaptive radio resource allocation in OFDM wireless networks with traffic prioritization," IEEE Transactions on wireless communications, vol.8, No.1, Jan. 2009.

[13] S. Niida, S. Uemura, and H. Nakamura, "Mobile services - User tolerance for waiting time-," IEEE Vehicular Technology Magazine, vol.5, No.3, pp. 61-67, Sep. 2010.

[14] S. Chen, J. Lukkien, R. Verhoeven, P. Vullers, and G. Petrovic, "Context-aware resource management for end-to-end QoS provision in service oriented applications," GLOBECOM workshops 2008.

[15] Y. Liang, X. Zhou, Z. Yu, H. Wang, and B. Guo, "A context-aware resource management framework for smart homes," Proceedings of the 5th International Conference on Ubiquitous Information Technologies and Applications (CUTE), 2010.

[16] M. Proebster, M. Kaschub, and S. Valentin, "Context-aware resource allocation to improve the quality of service of heterogeneous traffic," in proc. ICC 2011, June 2011.

[17] S. Niida, S. Uemura, and H. Nakamura, "A field study of waiting time filler delivery system," Mobile HCI 2011 (to appear).

 ISBN: 978-1-61208-167-0

Business Context Information Manager: application to Information Retrieval

Hamdi Chaker[1], Max Chevalier[2],
Chantal Soulé-Dupuy[1]
[1, 2]IRIT / SIG-D2S2u
[1]Université Toulouse 1 Capitole
[2]Université Paul Sabatier
Toulouse, France
{chaker, chevalier, soule}@irit.fr

André Tricot
Cognition, Langage, Langue, Ergonomie (CLLE)
Université de Toulouse Le Mirail
Toulouse, France
Andre.tricot@Toulouse.IUFM.fr

***Abstract*—Taking into account the context is important to improve the way systems provide relevant information to users. For this purpose, we introduce a *business context information manager* based on a novel and generic interpretation of the context. This manager takes into account various contextual dimensions and acts as an intermediary between information retrieval system and contextual information. The approach presented in this paper is suitable to any business context but in the paper we particularly illustrate it in Information Retrieval (IR) field. This approach relies on an original process (MES) that manages the various contextual dimensions to create a unique situation at a moment *t*. To this end, MES uses rules set which is the knowledge of the context manager. The situations will be used by a third part application (i.e., IR systems) for activity adaptation. Furthermore, an extracting process is also proposed to improve the context manager reliability over time and to facilitate its knowledge evolution. Finally, the proposed BCIM have been implemented to demonstrate the feasibility of our approach.**

Keywords-Context information management, Contextual Dimension, Situation, Contextual Information Retrieval, Business tasks, Information-related tasks.

I. INTRODUCTION

The aim of any Information Retrieval System (IRS) is to meet the needs of the user by providing relevant information. For this, the current trend is focusing on the user in order to serve him in the best way. This allows the systems to take into account the heterogeneity of users and the diversity of their needs. Therefore, the system maintains a representation of the user in what is commonly called a user profile. However, considering only the user profile to meet the user needs is not an answer. Indeed, every user's tasks of Information Retrieval (IR) may be performed quite differently depending on the context in which it was carried out. As a solution, the IRS gradually become "context-aware". This motivation is especially important in corporate settings (aeronautical, automotive, etc.) where tasks are critical and should be performed from specific information. In this context, the professional users (operators) require from IRS to provide them with accurate information necessary to perform their business task (defined and formalized as part of a business activity).

In this direction, we propose an approach to manage business context information enabling any IRS relying on it to have a realistic snapshot of the past and present context in order to perform finer adaptation of the information returned to the user.

The proposed Business Context Information Manager (BCIM) is based on an original interpretation of the context which assumes that the contextual information is dependent on each other. That is to say, the BCIM particularly focuses on the analysis of the interactions between contextual dimensions. This approach is intuitive and takes into account all kinds of contextual dimensions that can be modelled and valuated.

To model context and generate various situations, our approach is based on a set of rules and a specific process called MES. This process aims at contextualizing the various context dimensions on which the adjustment is desired. MES handles the adaptation of the different dimensions using these rules and creates a unique situation at a moment *t*. This situation is a stable interpretation of the context and all the past situations compose the context history. Another process introduced in this paper gains knowledge over time and then improves the context manager reliability. Indeed, the rules extraction process extracts new rules from past situations allowing MES to achieve a finer contextualization.

In this paper, we motivate our approach by giving the four main objectives of the BCIM towards the IRS before presenting the BCIM with a particular emphasis on MES process and the rule extracting process. The proposed BCIM is intended to the IR field. Specifically, the dimensions considered are the user, the business task and the environment (location, equipment, etc.). These business tasks need information to be performed by users.

Hereafter, the paper is organized as following. Section 2 presents background on context definitions and their use in the field of computer science and especially in IR field. Section 3 introduces the three contextual components of the IRS business context. Section 4 discusses the main purposes of the BCIM as well as its overall architecture. In this latter section, we also present the two main BCIM processes: the MES process and the rules extraction process after giving the various types of rules used in our approach.

II. RELATED WORK

Since we focus on IRS in this paper and since the general concept of context is broad, we concentrate in this section on the field of IR after giving some definitions of the context in the literature. We also discuss some limits of the IRS in

 ISBN: 978-1-61208-167-0

previous work. In the end of this section we give our interpretation of the notion of "context".

A. Context in the literature

The existence of numerous definitions comes from the multidisciplinary and rich nature of this notion. One of the broadest definitions of the context was proposed by Schilit and his colleagues. They claim that the important aspects of context are: where you are, who you are with and what resources are nearby [1]. The physical location of the operator situation is the central part of context proposed by Schilit and his collaborators. Pascoe [2] defines context as the subset of physical and conceptual states of interest to a particular entity.

One of the most accurate definitions is given by Dey [3] "Any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves". From this definition, the context is a set of situations and actions. Such situations change over time, it depends on the behaviours of users, on the applications and finally on environmental states at the moment when actions are performed.

Among the multitude of context definitions, it is important to introduce the work of Dourish [4] to have a clear vision on it. The author introduces taxonomy of contexts, according to which contexts can be classified into the representational and the interactional views. In the representational view, context is defined with a predefined set of observable attributes which does not change significantly over time. Thus, the representational view assumes that the contextual attributes are identifiable and known a priori. In contrast, the interactional view assumes that the user behaviour is induced by an underlying context, but that context itself is not necessarily observable. More interesting, Dourish [4] assumes that different types of actions may give rise to different types of relevant contexts. Thus for Dourish, there is a bidirectional relationship between activities and underlying contexts. In other words, contexts influence activities and also different activities giving rise to different contexts.

B. Context in IR

The issue of integrating the context in the IRS lies in the simple fact that a system cannot display the same result for two users retrieving information in two different contexts only because they have expressed the same query [5].

First, IRS can be improved by (1) modelling, (2) integrating, (3) using the context. Thus, the context can be used for example to improve the way people formulate their needs to the IRS and explore the returned information [6]. Traditionally, important contextual variables are included: user contexts (for example, its fields of interest in the short and long term, its habits, etc.); object contexts; tasks and social contexts where information needs arise.

There are several types of dimensions that can be integrated to a context in an IRS. The most important dimension is the user's profile, or more precisely the user's area of interest [7]. Context-aware IRS can take many other dimensions such as the nature of the task or the environment of the search to adapt the retrieval process [8]. Cool and Spink [9] distinguish four dimensions for the contextualization in the field of IR: information environment level, information seeking level, IR interaction level and query level.

At the present time, almost all work in this area are dependent on the explicit specification of the search goals, information related tasks and user intentions [10].

We go further to introduce information tasks and their relation with the business tasks. These tasks received particular attention in recent work to improve context-aware IRS. These studies concentrate on the users behaviour searching information and on the tasks that motivate his search process.

The relationship between business tasks and the information related tasks has been underlined by Byström and Hansen [11]. They consider the information related tasks carried via the IRS (information-seeking, information searching) as sub-tasks of business tasks. A relationship between the various information-related tasks has been also available. Recent work as those of Li and Belkin [12] offers a faceted approach to conceptualizing information related tasks in information seeking to explore the relationship between business tasks and the interactive behaviour of information access. Ingwersen and Järvelin [13] have another view of the context in the field of information retrieval. Their decomposition of the context is centred on the user achieving his business task related to information task(s). The information related task is always included in a business task which is itself the motivation of information search.

C. Limits of context-aware IRS

Before discussing the limits of previous work, it is important to point out that most of previous approaches agree on a common core that includes environment and human dimensions, but differs on the elements that must be included in the context [14].

One of the current IRS limits is that they do not have access to the business tasks and the possible conjunctions between all parts of the overall context [10, 11]. In addition, the specificity of IR in business context is to use the IRS to find the (missing) information necessary for business task achievement, i.e., the business task belongs to a tasks hierarchy; which is linked to other business tasks that have pre-conditions achievements, which make the classic IRS inadequate. Thus one of the current IRS limits is that they do not have access to this business tasks modelling.

Furthermore, previous work tried to understand how the user is doing his information task in his work environment. For these approaches, a better knowledge of the user allows adapting the search process to meet his needs [15]. However, the scope of previous work is limited because they do not converge: they are interested only in partial aspects of the user tasks or in specific business context.

Thus, the main objective of the business context information manager is to provide the IRS with realistic

 ISBN: 978-1-61208-167-0

situations. These saved situations from the context history can be used by IRS in two ways:

- *Present use*: context history used if it is relevant to current user need.
- *Future use*: context history used to predict user need or contextual element values.

D. Definitions

According to the interactional view of Dourish [4], as described at the end of the section II.A, the context influences activities and similarly different activities give rise to different contexts. In addition, some relevant work underlines the fact that the context is relative to something in particular: the context of an action, background interactions, etc. [16]. Thus, in respect of these views and the definition of context given by Dey [3], we define a situation as:

Definition *A situation is a stable interpretation (snapshot) of the context of an object at a specific time t.*

Every situation is characterized by a set of information (cf. Dey definition) which is organized in contextual dimensions to describe the various complex elements (i.e., user, system, task, environment, etc.) implied in the context. So, we define a contextual dimension as:

Definition *A contextual dimension composing the context of an object describes an external element that may have an impact on this object.*

This definition highlights that we limit the situation to elements which eventually have an impact on the object which we want to model the context. This has been decided to avoid information overload issue in the context representation. The choice of contextual dimensions is predefined and can vary from one application to another. We note that for our approach the contextual dimensions are not predefined a priori by the system. With regards to the user applications, the contextual dimensions and their contextual elements are selected in each situation.

Definition *Contextual elements are the leaf nodes composing the contextual dimensions trees.*

A contextual element is characterized by their name, their value, a Boolean that indicates if the element can evolve (i.e., for the contextual dimension "User", an element "name" could be considered as constant unlike to an element "tiredness" that can evolve through situations). The belief value (real number) allows the context manager to gradually measure the accuracy of the value. Indeed, some changes can be done due to the interactions between contextual dimensions. So it is important to give to the IRS the belief value relatively to the contextual element values.

In this paper we particularly focus on the dynamics of the interactions between the contextual information as underlined by Dourish [4]. Indeed, interactions exist between various contextual dimensions. For example, consider the following context for a specific information system composed of two contextual dimensions: user and task. The task dimension may have an interaction with the user dimension because this latter becomes stressed when he is performing it. That is to say, we cannot consider contextual dimensions independently within a situation. This makes reference to the **stability** property in the situation definition. Consequently to generate a new situation BCIM have to take into account all the interactions between contextual elements to ensure the situation stability. That is to say, no more interactions exist between contextual dimensions in this situation.

III. WHAT IS THE CONTEXT OF IRS?

We present in the following the three interdependent dimensions that compose the business context of an IRS. This triptych includes: users modelling, task modelling and environment modelling. As described in section II.B, these three dimensions are considered in IR field as the most important contextual dimensions.

A. The user

The *user dimension* corresponds to all contextual (long-term) factors related to users. It is based on a user model. Several elements can be considered as essential for our business context and thus include in the model of the user. These elements come from most of personalized information access work. This model has to be as general as possible in order to be used in different applications as proposed in [15].

B. The business task

The original dimension integrated to context-aware IRS concerns *business tasks*. It includes business task and its relations to information related tasks.

We show (Fig 1) an example of relation between one business task (and its sub-tasks), requiring information and two information processes allowing supplying the missing data necessary for its processing. The information process is not unique for a given task (i.e., T1.1). Actually the IRS has to select the most adapted information related task for the user in his environment.

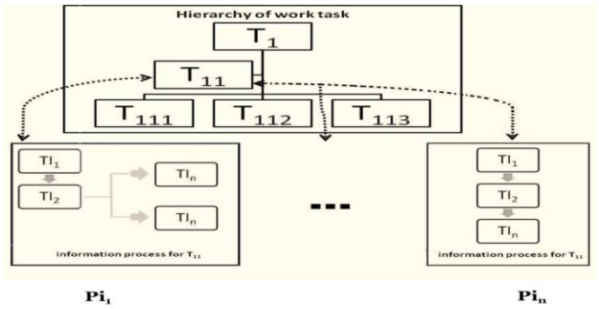


Figure 1. - Example of a hierarchical business task and its relation with information processes.

C. The environment

The *environment* is the dimension that aims at modelling all the environmental factors. The context can be interpreted as the environmental information in which the exploitations of the information take place.

IV. BUSINESS CONTEXT INFORMATION MANAGER

We propose a BCIM in order to reach main goals of context management for IRS. To achieve these goals, IRS

and BCIM exchange information like contextual information, actions done by end-user, etc.

In the following, we introduce the BCIM main purposes.

A. Purposes of the proposed BCIM

We identified four main purposes related to context information management that are important for business context-aware IRS.

1) Providing the contextual elements to IR

This first purpose is the most common one. Indeed, context-aware IRS should have access to contextual information (e.g., for systems adaptation). For this purpose, the BCIM enables the IRS to query contextual dimensions of the current situation. As contextual information can be "fuzzy" or approximated, every value given by the BCIM is associated to its belief level.

2) Checking situations validity

The second objective of our BCIM concerns the checking of the validity of any situation. A valid situation is a situation that actually satisfies the constraints and rules. Invalid situations are identified and the system alerted.

3) Anticipating actions

The third objective of BCIM is to anticipate actions and to recommend them automatically to the IRS for adaptation purpose for instance. When generating a situation, the BCIM identifies similar situations in order to identify actions that have been the most used in such situation. As a consequence, for a given situation, BCIM can recommend to IRS as well as to end-users a specific way to achieve their goals (through recommended actions).

4) Addressing the lack of information in the context

The fourth objective of the BCIM is the completion or the adaptation of situation content according to the knowledge related to interactions between contextual dimensions. Indeed, sometimes some contextual information is missing or lacks of accuracy. As a solution, the BCIM exploits the knowledge it can extract from past situations to complete such contextual information.

B. The knowledge exploited to achieve the BCIM goals

In this paper, we present three types of rules used by our BCIM. Note that the two first kinds of rules are provided a priori by domain experts.

- **Legal business rules**: they correspond to rules whose violation is strictly prohibited by legislation (e.g., law of the company). It is necessary to comply with these rules.
- **Business rules:** they are provided by experts of the domain to enhance activities performance or to establish procedures or processes. It is not required to comply with these rules. These rules can evolve according to users' real activity.
- **Inferred rules:** such rules are extracted via a rules extraction process which is detailed in the section IV.C.2.

Every rule is given as an implication "X implies Y", noted **X => Y**, where X is a conjunction of some contextual elements from user, task or environment. Y is a single contextual element that is not present in X. Every rule in rules set has also a priority which is computed by the BCIM or given by the domain experts. A rule is applied if and only if every contextual element in X is satisfied.

C. Overall architecture of BCIM

Thanks to these elements, now we introduce the overall architecture of our context manager with particular regard on MES and the extracting rules process (Fig 2).

The manager we propose contextualizes IRS. Indeed, the situations proposed by our context manager are the entry point of IRS. In other words, our approach allows the IRS to obtain the most realistic picture of contextual information in order to better satisfy / match the needs of the user.

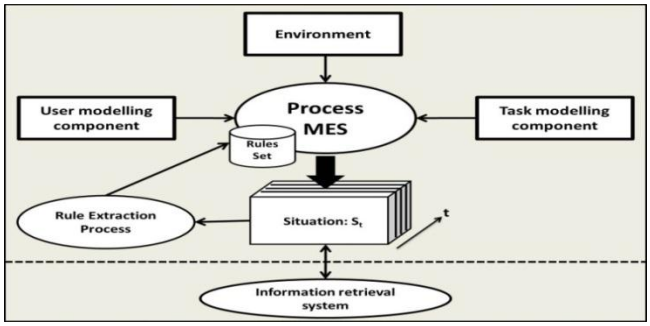


Figure 2. The overall architecture of the BCIM

1) MES process

During initialization stage, when BCIM is creating the first situation of a particular activity, the IRS relying on our manager provides MES with values of all the contextual elements of each dimension. Then MES checks the validity of the situation (according to the second purpose of the BCIM) and alerts the IRS in case of an invalid one. This may happen in two cases:

- The contextual elements from various contextual dimensions cannot interact into the situation because of the breach of one or more legal business rules. We can give the example of the aircraft maintenance field: technician with "class A" skills cannot perform intensive tasks or work on the tarmac.
- In addition to contextual elements, actions sent by the IRS may also violate such rules. We can give the example of the access to an electronic document that can be read only by specific categories of users.

We notice that the BCIM (through the MES process) stores the invalid situations to send feedback to the system experts. This can be helpful to improve the BCIM by updating the set of legal business rules.

Thereafter, when the IRS modifies one or more initial contextual dimensions (change in contextual element value), MES uses the rules set to adapt all other contextual elements to build an up-to-date situation. In other words, the goal of MES in this case is to assess the impact of the changed element value on the rest of the context and above all the missing contextual elements values. Thus, MES keeps the

 ISBN: 978-1-61208-167-0

given values and it predicts missing ones using the rules set (refresh the other values in the new situation).

We note that MES does not perform any adaptation when the situation is the most realistic interpretation of the context, i.e., all the values are correct and given by the system sensors.

Furthermore, it is important to give some clarification regarding the fourth goal of the manager. The lack of contextual information can occur in two cases:

- Upon initialization stage, there may be missing contextual information caused, for instance, by a failure in one sensor. To infer this missing information and to create the situation, MES applies the different rules in the following order: business rules and then inferred rules.
- Between two linked situations (situation at a moment *t* and situation at a moment *t-1*). To create a new situation (at a moment *t*), MES gets contextual information from the previous situation (at a moment *t-1*) in addition to new contextual element values supplied by the IRS. MES therefore can use all illegal rules types to adapt the contextual elements.

a) The adaptation cycle of contextual dimensions

For the effectiveness of any IRS based on our context manager, it is essential for all contextual dimensions to be adapted and confronted to each other (i.e., contextualized). The goal is to characterize the best possible situation. To achieve the adaptation of one dimension, depending on the others, the MES process uses the three types of rules to describe the various transformations to be applied. As detailed in the previous section, the specific type of rules used by MES widely depends on the precise purpose to be achieved.

Hereafter we present a simplified example of the three contextual dimensions in IR field (Fig 3). We note that these contextual elements are provided by the IRS to the BCIM. Thus to generate a new situation, the BCIM (thanks to the MES process) performs a new adaptation cycle.

Figure 3. Example of the three contextual dimensions

Suppose the following set of five rules: R2, R3 and R4 are business rules provided by the domain expert, R1 and R5 are inferred rules given by the extracting process. We remind that these rules are sorted according to their own priorities and they are applied if and only if their conditions are satisfied.

R1: User.Experience level≤4 AND Environment.Location=Office
➔ Task.required time=50min (priority = 0.9).
R2: Task.required time>40min AND Environment.Temperature>25°C
➔ User.Emotional state=stressed (priority = 0.8).
R3: User.Emotional state=stressed AND Task.Objective=read an instruction
➔ Environment.Background sound=OFF (priority = 0.6).
R4: User.Experience level>8 AND Environment.Location=Home
➔ Task.required time=5min (priority = 0.5).
R5: User.Emotional state=stressed AND Environment.Background sound=OFF
➔ Task.required time=70min (priority = 0.5).

By taking this example, we want to explain the adaptation cycle and how MES could adapt the various contextual dimensions. In our case, at the beginning of the adaptation cycle only rule R1 can be applied because it is the only rule whose condition is satisfied. As a consequence of this adaptation, the time required in the adapted task dimension (*Task'*) evolves; the value becomes 50 minutes instead of 10 min (see Fig 4).

Figure 4. The adaptation of the Task dimension

b) Situation stability

The adaptation cycle respects our context definition and its originality. Indeed, the situation stability gives rise to the dynamic interaction between contextual elements we wanted to reach. Every dimension is adapted by the use of all other dimensions. We therefore emphasize that the construction of a situation is a recursive process. Indeed, we cannot consider each dimension of a context separately. Thus, these rules are applied successively on the dimensions until a stability point of the context at a moment *t* is found; that is what we call situation. A stable situation can arise from large number of iterations (the stability process).

Taking the previous example (Fig 3), only R1 was applied at the end of the first iteration of the adaptation cycle. Therefore, to obtain situation stability, MES performs a second iteration and the new task (*task':* the adapted one) substitutes the initial task in the context dimensions set. As a consequence of *task'* elements values, R2 condition is satisfied. Thus, MES continues the stability process and adapts the user dimension using R2. Therefore, this adaptation infers knowledge on the emotional state contextual element of the user by altering the initial model and the emotional state of the user becomes stressed. In the same way, the new adapted contextual dimension (*user')* will impact on the environment since this change satisfies the R3 condition. In other words, R3 will be applied by our stability process and a new environment dimension feature is inferred by the MES process. It is important to note that R2 and R3 could not have been applied if only the initial contextual dimensions (Fig 3) had been considered without incorporating the adapted ones in the adaptation cycle.

As presented before, the three contextual dimensions are adapted at the end of the third iteration of the adaptation cycle. At this point MES will not stop the adaptation cycle and a fourth iteration is performed. Consequently, R5 can be applied and the task dimension is adapted again (time required becomes 70 minutes). Finally, at the end of the adaptation cycle a stability point is reached (no more rules can be applied). MES provides the IRS with the three

 ISBN: 978-1-61208-167-0

adapted contextual dimensions (see Fig 5) which are considered as the realistic situation.

Figure 5. The adapted contextual dimensions composing the situation

2) Rules Extraction Process

The BCIM must be able to evolve over time. In other words, the set of rules used by the MES process should grow to enhance contextual information accuracy. For this purpose, we propose a novel rules extraction process based on past situations. This process concentrates only on valid situations because they contain information about what happens in practice. Furthermore, the impact of the dynamic interaction of various contextual elements during the adaptation itself can cause a change in activity or in the initial models. Learning from past situations can give to our BCIM a realistic view of what actually happens during an activity. The system becomes then more and more efficient and the information more and more relevant as well.

For the BCIM, the rules extraction process is based on the classical association rules extraction methods [17]. These methods aim at finding interesting associations and/or correlation relationships among large sets of data items. Association rules show the value of items that occur frequently together in a given dataset. In our case, these items are the contextual elements and our dataset is the set of past situations. Thus, via this process, the context manager will enrich its rules set. Consequently, this will improve the quality of the adaptation of the contextual dimensions. We note that inferred rules are subsequently ranked by the BCIM according to their priority which is computed by rules extraction process using, for example, confidence or support.

3) Implementation

The proposed BCIM have been implemented as a Java server. The different contextual dimensions are stored in XML format. This choice is motivated to improve the possibility to describe various contextual dimensions (e.g., CTT approach: "ConcurTaskTrees" for hierarchical task modelling can generate XML document [18]). To extract and manage association rules, the prototype uses Weka API [19] and specifically its *a priori* algorithm.

V. CONCLUSION

In this paper, we present a BCIM for IRS business context based on three contextual dimensions: business task, user and environment. The manager is based on a novel interpretation of the context. This interpretation is original because it focuses on the dynamics between all these contextual dimensions rather than how information is represented. We introduced the MES process that manages the three contextual dimensions using the context manager knowledge related to interactions between contextual dimensions. The BCIM provides the IRS with realistic snapshot of the user activity. This snapshot is the current situation that will be used for any purpose by IRS. We have also shown that the three kinds of rules are useful to model knowledge necessary to manage IRS context. Finally, we presented the rules extraction process that improves the BCIM over time by making it gain new knowledge from past situations. As a matter of future work we plan to evaluate the BCIM using real situations in concrete business context. Then, we can compare other rules extraction methods.

REFERENCES

[1] B. Schilit, N. Adams, and R. Want, « Context-aware computing applications », in *First Workshop on Mobile Computing Systems and Applications*, 1994, pp. 85-90.

[2] J. Pascoe, « Adding generic contextual capabilities to wearable computers », in *Proceedings of the 2nd International Symposium on Wearable Computers*, 1998, pp. 92-99.

[3] A. K. Dey, « Understanding and Using Context », *Personal Ubiquitous Comput.*, vol. 5, n°. 1, pp. 4-7, 2001.

[4] P. Dourish, « What we talk about when we talk about context », *Personal Ubiquitous Comput.*, vol. 8, n°. 1, pp. 19-30, 2004.

[5] J. D. Johnson, « On contexts of information seeking », *Information Processing & Management*, vol. 39, n°. 5, pp. 735-760, 2003.

[6] G. Kumaran and J. Allan, « Adapting information retrieval systems to user queries », *Information Processing & Management*, vol. 44, n°. 6, pp. 1838-1862, 2008.

[7] J. Bai and J.-Y. Nie, « Adapting information retrieval to query contexts », *Information Processing & Management*, vol. 44, n°. 6, pp. 1901–1922, 2008.

[8] N. Stojanovic, « On the role of a user's knowledge gap in an information retrieval process », *Proceedings of the 3rd international conference on Knowledge capture*, pp. 83–90, 2005.

[9] C. Cool and A. Spink, « Issues of context in information retrieval (IR): an introduction to the special issue », *Information Processing & Management*, vol. 38, n°. 5, pp. 605-611, 2002.

[10] N. J. Belkin, « Some(what) grand challenges for information retrieval », *SIGIR Forum*, vol. 42, n°. 1, p. 47-54, 2008.

[11] K. Byström and P. Hansen, « Conceptual framework for tasks in information studies: Book Reviews », *J. Am. Soc. Inf. Sci. Technol.*, vol. 56, n°. 10, pp. 1050–1061, 2005.

[12] Y. Li and N. J. Belkin, « A faceted approach to conceptualizing tasks in information seeking », *Information Processing & Management*, vol. 44, n°. 6, pp. 1822–1837, 2008.

[13] P. Ingwersen and K. Järvelin, *The Turn: Integration of Information Seeking and Retrieval in Context (The Information Retrieval Series)*. Springer-Verlag New York, Inc., 2005.

[14] P. Brusilovsky and E. Millán, « User Models for Adaptive Hypermedia and Adaptive Educational Systems », in *The Adaptive Web*, 2007, pp. 3-53.

[15] M. Chevalier, C. Julien, C. Soulé-Dupuy, and N. Vallès-Parlangeau, « Personalized Information Access Through Flexible and Interoperable Profiles », in the International Workshop on Personalized Access to Web Information, France, 2007, pp. 374-385.

[16] J. C. Pomerol and P. Brézillon, « About some relationships between Knowledge and Context », *Modeling and Using Context*, pp. 461–464, 2001.

[17] R. Agrawal, T. Imielinski, and A. Swami, « Mining Association Rules between Sets of Items in Large Databases », in *International Conference on Management of Data*, 1993, pp. 207-216.

[18] F. Paterno, « Model-based design of interactive applications », *intelligence*, vol. 11, n°. 4, pp. 26-38, 2000.

[19] http://www.cs.waikato.ac.nz/ml/weka/, july 18 2011.

 ISBN: 978-1-61208-167-0

Privacy and Distributed Tactical Operations Evaluation

Dennis Andersson
Information Systems
Swedish Defence Research Agency
Linköping, Sweden
dennis.andersson@foi.se

Abstract— **In this study, thoughts on ethics of workplace monitoring are being applied to the very special domain of evaluations of tactical operations, such as military or crisis management exercises or operations. I try to find out if there are differences in the way we should regard workplace monitoring when it comes to this domain compared to standard workplaces such as offices, since the purpose of the surveillance is not to enforce discipline, but to evaluate the organizations' ability to conduct a tactical operation. The study focuses on issues such as privacy and informed consent and the main purpose of the investigation is to structure a consistent ethical standpoint when it comes to operations' evaluation by making parallels to related theories that I found correct and applicable. I conclude that is indeed reasonable to place other demands on crisis management workers than we would do on other work forces, and that it should therefore be easier to motivate workplace monitoring for the purpose of evaluating distributed tactical operations. I argue however, just as Miller does regarding police work, that upholding privacy can be a real problem when crisis management personnel are exposed to monitoring, even though it is intended for evaluation.**

Keywords - privacy, workplace surveillance, after-action review, crisis management

I. INTRODUCTION

The strive to develop and increase efficiency is, and has always been, an important force in society. Any organizational development assumes some sort of evaluation of the current state, to validate or verify the organizations' processes. Operational organizations such as the Armed forces, the Police and the Rescue services are not exceptions to this rule. Hence, evaluating their operations can be fruitful to yield understanding of how the organizations function, and thus a lot of effort is currently being put into developing methods and technology for such evaluations. In this sense, technology development is a door opener enabling a new spectrum of analyses; using video cameras, audio recorders, radar, position trackers and other sensor systems. From an ethical point of view, we need to investigate how those being evaluated react to the inherent analysis of their actions, which from an ethical point of view can be compared to the commonly discussed topic of workplace surveillance.

Workplace surveillance has been around since the early days of industrialization and is a powerful tool for the employer to ensure that the employees are performing the work they are hired to do. As technology progresses it becomes easier for the employer to increase the level of surveillance on the employees at the cost of an increased risk that they will experience the surveillance as a violation of their privacy. Kizza and Ssanyu [1] states that employees can get a reduced self esteem and become less creative by intrusive surveillance. In their article they present a number of arguments for and against workplace surveillance. I will return to the most relevant of their arguments later in this paper.

II. OBJECTIVE AND READING DIRECTIONS

In this paper I leverage on the work done on workplace surveillance by Kizza and Ssanyu [1] among others to direct the same questions to the question of monitoring crisis management personnel, where the purpose of the surveillance is not to enforce discipline, but to evaluate the organizations' ability to execute a tactical operation. The main purpose of this study is to structure ethical guidelines to consider for operations' evaluation by making parallels to related theories that I found correct and applicable.

The questions dealt with here would benefit from a larger empirical study of crisis management work, but that lies beyond the scope of this paper. Instead the study is focused on existing literature on nearby topics to yield some insight into the questions that need to be asked when conducting such an empirical study. To bring the findings in [1] closer to the crisis management domain, I take help from Miller [2] and his reasoning around the problems related to surveillance of police officers in their daily work. He argues that there really is a difference in what kind of surveillance you can expect a policeman to accept compared to for instance an office worker. The main contribution of this paper is the study of what privacy issues may arise when monitoring personnel in distributed tactical operations. The case studied here differs from the typical case of workplace surveillance, since the surveillance serves the purpose of validating and verifying processes and are therefore typically only in place for a limited time period, which was not the case in for instance Miller's much quoted study.

I argue, just as Miller does regarding police work [2], that upholding privacy can be a real problem when crisis management personnel are exposed to monitoring for the purpose of operation evaluation. In the following chapters I present my arguments before I finally structure my stand on evaluation of tactical operations. I start by defining some key

 ISBN: 978-1-61208-167-0

notions and describing the context around which my reasoning revolves.

III. Evaluating Distributed Tactical Operations Using After-Action Review

Evaluation of distributed tactical operations is here used to denominate the systematic practice of evaluating a tactical operation or exercise spanning across multiple locations [3, 4]. Originally, the approach was designed to evaluate military exercises; however it has also been successfully applied in many civilian domains, such as first responders in crisis management operations. In large and complex operations with many organizations, the after-action review (AAR) [5, 6, 7] can be more or less independently executed for each unit in the training audience. While these small group session AARs remain important, large-scale shared AARs receive more and more attention. In these overarching AARs, the evaluation focuses more on strategic command and control (C2) than the regular AARs that typically deal with issues on the operative level. Exactly where to put the efforts will depend on what kind of issues you want to highlight for the training audience.

The concept of AARs was coined by the U.S. Army who, like many other organizations, conducted review sessions after each exercise and operation. It was formalized and labeled AAR in [5]. First responders and crisis management organizations often use the same methodology, sometimes labeling it hot wash or debriefing. There is nothing in the theory itself that defines how to capture data for preparing an AAR, but more often than not, the main source for data capture and input to the AAR process are human observers/trainers (OT). An OT can offer a subjective view of what happened during the operation, and can use his/her expertise to pinpoint interesting events for AAR discussion. However, technology is quickly gaining ground in this area as a complement to the human observers. Audio and video recordings combined with system logs and other sensory data can provide an undisputable ground truth that the AAR facilitator can use to provide a baseline for the audience to review and discuss. As technology advances, the quality and the quantity of this data increases as well, enabling more and more accurate and detailed reconstruction of the events.

IV. Reconstruction and Exploration as an Approach for Conducting AAR

Reconstruction and Exploration (R&E) [3, 4] is a formalized approach to, among other things, support an AAR after a complex chain of events, such as a military operation or a crisis management operation. The approach assumes both human observers and technical registration to collect information on the chain of events, very much like the practice that AAR facilitators are moving towards. The aim of the data collection is to gather enough data to enable reconstruction of the operations as a time-synchronized visual multimedia model. The model can be used to find system- and organizational problems and identify needs for improvement. However, it is important to note that not only negative feedback is captured and reported, it is commonly recognized that positive feedback is equally important to provide during the AAR session.

Some of the most common data sources used for R&E are observer reports (notes), video surveillance systems, handheld cameras, microphones, wiretapping, screen capture systems and GPS devices. Exactly what data sources are used will depend on what questions need to be answered or what hypotheses are tested during the exercise/operation. In some scenarios there are no predefined questions or hypotheses, in which case the AAR team will typically try to collect a data set that is as comprehensive as possible to be able to answer any questions that may arise during or after the operation. This all makes R&E-assisted AAR a very powerful and flexible way of evaluating exercises and operations, especially distributed ones where an OT can have a hard time getting a birds-eye view of the scenario until being presented the data in the exploration phase.

Kizza and Ssanyu [1] describes workplace monitoring as a dominance or power between workers and employers, where the purpose is to:

- Increase productivity,
- protect against theft,
- protect against espionage,
- performance review of employees,
- prevent harassment,
- find missing data,
- find illegal software or
- prevent personal use of company resources.

Data capture for the sake of R&E partly adheres to the fourth bullet, although the purpose of an AAR is usually to assess the performance of a process or an organization as opposed to an individual as was the case in [1]. This difference may be crucial to the training audience's acceptance of monitoring, but it becomes apparent that R&E could in theory be used for all of these 8 purposes too, which would incur reduced trust, both for the system and the OTs.

V. Crisis Management Work

A crisis is sometimes defined as something that threatens basic functionality and values of society or individuals. That is a too broad of a definition for this work, as it spans from natural disasters to personal tragedy. Instead I will only use the word crisis as the type of extraordinary event, disaster if you will, that affect society as a whole. More specifically I will focus on events that require interagency cooperation. Some of the most recent examples from the Swedish society include the 'Gudrun' storm in 2005 [8], the Indian Ocean tsunami in 2004 [9], the discotheque fire in Gothenburg 1998 [10] and the M/S Estonia disaster in 1994 [11]. The evaluation of this kind of events are often handled as special investigations by appointed authorities, in Sweden typically by the Accident Investigation Board, who tries to analyze what happened and clarify whether there were any mistakes or procedural errors that need to be fixed. Considering that you can never fully prepare for an extraordinary event, it is in reality impossible to guarantee that you will be able to recover all the data you need for this kind of analysis. When it comes to exercises on the other hand, the course of events

 ISBN: 978-1-61208-167-0

can be controlled, and collecting the right amount of data at the right time is a matter of thorough preparation. For this reason, R&E is best suited for exercises and well-planned operations as opposed to the chaotic environment that first responders typically are exposed to in a major disaster, and where R&E may be harder to apply.

Typical for this type of extraordinary events is that they put routines to the test and may even become impossible to apply. Ad hoc workflows may have to be created as well as spontaneous command and control structures that will help in dealing with the situation. To assess performance in this scenario is difficult as you cannot always foresee what the processes will actually be. R&E gives some support in this process as it offers a rich data set and is flexible in the way data is being used and analyzed.

VI. PRIVACY

There is vast number of known and used definitions of privacy, one of which was formulated by Warren and Brandeis in 1890: 'the right to be let alone' [12]. This definition is still in use, but not very suitable for privacy in the professional life, instead I will rely on the definition that Aiello and Klob used in their publications on workplace surveillance [13]: "Privacy is the ability for an individual to control the use of their own personal data, wherever it might be recorded".

By the definition above, every human has the right to control any information about them and to avoid being seen. The immediate consequence of that is that all non-controlled surveillance and monitoring must be regarded as an infringement on privacy which of course is a problem. The keyword in this definition is 'control' which calls for further investigation and interpretation.

VII. PRIVACY INFRINGEMENT ISSUES

In the following section I will show that infringement is a real problem in monitoring crisis management personnel for the purpose of performance assessment, and then discuss arguments for and against acceptance this infringement, and lastly present and justify my personal opinion on the matter.

A. Is there a problem at all?

Kizza and Ssanyu [1] discusses ethics in technical workplace monitoring. The scene for a crisis management operation is a indeed a workplace, and there are many similarities between R&E monitoring and the methods and techniques that they mention; such as wiretapping, screen capturing, keyboard input logs, computer network surveillance, video recording, e-mail forwarding, etc. All of these technical monitoring solutions can be very useful in R&E, all depending on what aspects of work the OTs need to review. Hence, the arguments that [1] are using are worth considering for R&E.

The fundamental conflict of values concerning workplace monitoring is about the employees' right of privacy versus the employer's right to ensure that he/she gets value for his/her investments. That workplace monitoring does exist today is well known, and society seems to still be functioning, so maybe it is not a real problem after all? Kizza and Ssanyu [1] points at nine negative consequences of the workplace monitoring:

- Lack of trust between workers, supervisors and management,
- stress and anxiety,
- repetitive strain injuries because of refraining from taking breaks,
- lack of individual creativity,
- reduced or no peer social support,
- lack of self-esteem,
- worker alienation,
- lack of communication and
- psychological effects.

Some of these effects are the result of a lasting monitoring of an individual, and I do not believe that these are directly applicable to the domain of this study. However, it is reasonable to suspect that at least consequences 2, 4, 5, 6 and 8 exist also in the crisis management domain. The latter of them is something that can clearly be noticed during exercises as training audience sometimes turn off monitoring equipment to allow them to speak freely. This implies, not surprisingly, that monitoring does infringe on their privacy.

Palm [14] states, with reference to Alpert [15], that new technology has enabled employers to shift performance monitoring to target individuals as opposed to teams as was the case earlier. I second that opinion, and I see the same tendencies in the AAR domain, i.e. more and more technology-based solutions are implemented; making it easier to assess individuals rather than teams. As gadgets are easy to reproduce and relatively cheap, it does not always occur to the OTs that there may be a reason to minimize the amount of recording equipment instead of just adding more gadgets. The way to remedy this is to spend more time preparing the setup by thorough modeling and instrumentation planning in the initial phases of R&E, to carefully decide what recording is necessary and what is not. If this step is not properly managed, there is an apparent risk that the training audience gets a lower trust in the monitoring and develops a negative attitude towards it. However, with reference to Merz Smith [16], Palm argues that it is not just employers that benefit from this type of monitoring [14]. She continues to explain that employees can regard it as a positive experience that their hard work is being noted and that 'leeches' will have a harder time getting away with their laziness – especially since this type of monitoring is more objective than having a person watching over your shoulder and possibly favoring or discriminating among the employees. In the same manner, first responders and crisis management workers could benefit from monitoring, especially live operations, since they get an means to prove that they acted correct based on the information available to them at the time, and thus avoid criticism from the "all-seeing, never-knowing" public. This reasoning is something that Miller [2], among others, uses and I will return to it later in this paper.

Palm [14] especially mentions four risks of continuous and systematic collection of personal data:

- Unavoidability,

 ISBN: 978-1-61208-167-0

- continuity,
- dependency and
- identifiability.

By unavoidability she means that as an employee you have little or no say on what kind of data is collected about you, other than by changing jobs. By continuity she refers to the problems of continuous monitoring, which can have consequences for your privacy. The dependency issue she mentions is about the employee's dependency towards to employer and that there is therefore an asymmetric power relationship between them. The last bullet, identifiability, is negative according to Palm since it makes it easy to combine different data to find patterns and profiles that the monitoring system was not originally designed to do. She concludes that at the workplace you are more vulnerable to privacy violations than elsewhere. Whether her remarks are valid in the domain studied here is not easy to settle. For instance, it is obviously easy (and recommended) to give every member of the training audience a choice to accept monitoring and be part of the R&E evaluation, or to stay out, which should directly cross out the first bullet on her list. However, what is not clear is what will be the consequence of staying out. Is this person going to be replaced? Is the exercise going to continue as planned, with an altered instrumentation plan? Will the declining individual miss out on valuable training experience? If so, will that reduce his/her ability to operate in a crisis management operation? Ultimately, can the consequence become failure to save lives because of inadequate knowledge? In effect the unavoidability may still be a problem then.

The aspect of continuity that Palm mentions is probably not as relevant in this study, since data collection is only being done during exercises, or possibly on some live operations. There is no reason to continue the collection during regular duty, at least not for the purpose handled in this paper. The issue of dependency however may face the exact same problems as in her study since there is often a well-defined hierarchy of command in these organizations where the same power-issues arise. Identifiability can also become a problem for R&E since typically much of the data streams are associated with individuals, making it very easy to deduct personal information that the system was not intended for. To avoid this there are techniques to de-identify persons, but that can potentially cause problems for the OTs as they need to know who was responsible for decisions and actions to interview them on their thoughts at that time.

A contractarian would be able to claim that Palm's dependency relationship between employer and employee is in fact a contract where the employee gives up some fundamental rights to privacy by accepting the job offer, especially so in the public sector since tax payers have a reasonable right to demand that their tax money is used for greater good. Since both parties have agreed to this contract, the employer should then be entitled to perform this monitoring according to the contractarian. To further strengthen the argument, some employers are adding monitoring clauses into the employment contract to clarify that his rights trump the individual's right of privacy. By signing the contract the employee can be considered as having given consent to the monitoring and therefore there is not an ethical issue at hand. To counter this, you could argue that the employee in reality has no good options, since by refusing to sign he or she would be unemployed and have no income. For instance, many philosophers compare this to voluntary slavery, which according to the contractarian would not be a problem, while someone with a broader perspective would argue that there may be problems with information or other issues that makes the contractee not understand the consequences of what is being agreed. In the field of medicine the notion of *informed consent* is often used. Malek [17] defines it as 'voluntary consent based on adequate understanding of relevant facts'. She mentions five important parts of informed consent:

- That the subject is given all information,
- that the subject understands all information,
- that the subject is able and allowed to make a choice,
- that the subject makes the choice without involvement of a third party and
- that the subject actually gives consent.

Within the area of medicine, this form of consent is necessary to conduct certain procedures. Clarke [18] states, and I concur, that the same requirements should be applicable to infringements of privacy such as through monitoring. He points at several actual problems within this area, such as installation of new surveillance equipment without explicit consent from employees. For the sake of R&E evaluation, this does not necessarily impose a problem as the equipment can be setup temporarily and that the training audience can easily be informed of all the data collection that will take place.

Of greater relevance to this study is the notion of *continuous informed consent* that [18] describes as extra complicated since the subjects may find it difficult to grasp in what way the surveillance equipment will be used in the future. E-mail forwarding on the workplace, for instance, can be used to counter industrial espionage. Although this can itself be very controversial and sensitive to some, it is not difficult to imagine that some employees accept this privacy infringement and agree to setting up the system. When employers use the system to create detailed analyses of their employees friends and relationships to find persons at risk of being targeted by spies, it all of sudden becomes a lot more violating to privacy and it is not likely that the subjects would agree anymore. Therefore a mechanism is needed that allows informed consent to be revoked. This applies also to evaluation of distributed tactical operations since combining several data sources can enable detailed profiling of individuals and teams that was not obvious at the start. It is clearly relevant to ask whether the subjects' consent can be regarded as informed according to the definition in [17] when the objective of data capture may be unclear even for the OTs at the time when participants give their consent.

It would be very valuable to conduct empirical studies to decide how crisis management workers relate to workplace monitoring and give them a chance to give an informed consent. Such studies have unfortunately not been conducted for this particular paper, instead I will look at three empirical

 ISBN: 978-1-61208-167-0

studies [19] to continue my reasoning. They interviewed employees and students in Ireland and Great Britain to, somewhat surprisingly, conclude that employees do not regard workplace monitoring as problematic. Based on that finding, they question whether there is any point in discussing the ethics in it. A more detailed review of their studies shows that there seem to be quite a few interviewees that actually do consider monitoring as a problem, although the majority does not. This is a result that I find less surprising as the level of privacy infringement one can allow before feeling violated is highly personal. To state, as the authors do, that monitoring is then not problematic is to neglect that portion of the population that does, and I would say that their conclusion that employees do not consider monitoring a problem is therefore greatly exaggerated. They do, however, extend their reasoning and argue that surveillance reduces self-esteem and creativity among the subjects, just as Kizza and Ssanyu reported [1]. According to [19] this can happen without the subjects even realizing it, which makes the problem even more complicated, and again we have to revert to theories on informed consent. To count as an informed consent, the subjects need to understand what monitoring exists and how it affects to them, which according to [19] is not always the case.

First response and crisis management work differ from the kind of office work studied in [19] in the sense that workers are part of a process that fills an important role in society safety that we as tax payers and citizens rely upon and consider ourselves entitled to demand. Miller [2] makes a similar statement when he focuses on monitoring of police officers in their daily work. He notes that there are both differences and similarities compared to office workplace monitoring. He describes privacy as a morale right that all humans have, regarding control over information on themselves and how they are seen by others, sometimes referred to as the 'private sphere'. He argues that no matter who you are and in what situation you are in, the right of privacy always applies, and as such there is a problem in monitoring police officers since they in fact lose control over who sees them and how. From his reasoning we can deduct that workplace monitoring is an infringement on privacy regardless of workplace, and therefore also for the purpose of evaluating distributed tactical operations. My own conclusion is that there is a clear problem with workplace monitoring that does apply to exercises and operations of a distributed tactical character, such as the ones mentioned earlier. The problem lies in a violation of privacy of the training audience and it must be weighed against the positive effects that the R&E evaluation gives. How personnel reacts to this infringement on privacy can differ a lot, but I also note, with respect to [19], that problems can exist that the subjects are not aware of or has given consent to, since the effects can be subconscious, which according to [17] then negates the consent.

B. Can we demand that crisis management personnel accept an infringement on their privacy?

As argued above, there seems to be an infringement on crisis management personnel's privacy when being evaluated using AARs, and we must be able to motivate that this is an acceptable cost if we as a society are going to accept this infringement. In this section I will compare some of the pros and cons of monitoring, and try to relate that to the infringed privacy to establish a consistent view on R&E monitoring.

As [2] states, society as a whole benefits from a well functioning police, in the same manner it benefits from not only having well functioning crisis management organizations, but also just knowing that it works well can have a calming and positive influence on society. This means that a utilitarian could argue that privacy infringement on crisis management workers is accepted to create a better society. Ross' pluralism [20] tells us that there may more to the story and implies a paradox here, as we will have two duties facing each other: the duty of beneficence vs. the duty of non-maleficence. Which duty is our prima facie in this case is not obvious in the pluralistic deontology of Ross. An interesting recent such scenario, non-related to monitoring, is that of the triple disaster in Japan causing a nuclear crisis in Fukushima; any worker approaching the reactor faced an obvious risk of being exposed to lethal doses of radiation, not to mention risks of explosion and collapsing buildings. How can anyone be asked to go to work during such conditions? Meanwhile, society faced the risk of meltdown and an even larger calamity.

The utilitarian reasoning would be to put the duty of society's best first, which I can partly sympathize with. It is however, as so often in ethics, a trade-off and we have to be careful in our reasoning and not forget that sometimes creativity may suffer. In a crisis situation where routines and resources are not enough, individual creativity is often what drives the work. If we in our strive to evaluate our societal functions render them inefficient, it may mean that fewer lives can be saved at the next disaster, which will then be the price we pay to feel safer; a very disturbing and contradictive thought in itself. A consequentialist would of course argue that this is therefore the wrong path. Although the consequences here are stretched to the extreme, I argue that there are risks both in monitoring too much and too little.

A very important difference that Miller [2] mentions is that the police are expected to serve society and that they therefore are prepared to accept a higher degree of monitoring and a reduced privacy, and that this is all well-known to them when they apply for the job. He also points at pros in monitoring where police officers can use audio and video recording to prove that they acted correctly when being questioned after severe incidents.

The type of monitoring that Miller [2] deals with, is more or less constant during daily work. When it comes to evaluating distributed tactical operations, it is always a matter of well planned exercises that are out of normal work. This has two major implications for the reasoning in this study. Firstly, the infringement on privacy is temporary and thereby easier to accept. Secondly, it is not at all safe to assume that the workers are used to this type of monitoring and some may react different than those who are used to it. This can result in the workers being so intimidated that their creativity and performance become dramatically reduced.

 ISBN: 978-1-61208-167-0

The arguments in [2] are partly applicable to the crisis management domain. It is clear that the workers have a greater acceptance to monitoring than would employees at an ordinary workplace, especially in the distributed tactical operation scenarios that this study deals with, as the monitoring is temporary. However, even though I lack evidence thereof, I can still imagine that the monitoring can affect the training audience to the extent that the evaluation becomes counter-effective. To minimize the risk of that, it is vital to inform every member of the audience of the benefits of evaluation and clarify that AARs are all about generating feedback to the team and that the assessment is primarily on team level, not personal.

VIII. CONCLUSION

Evaluation of distributed tactical operations is an important tool to verify and validate crisis management work and refine it. Technology advances quickly and generates more sophisticated tools to analyze the work. In my opinion, society benefits from going forward with this type of evaluations, but we should be careful and aware of the processes that are activated at the training audience. There are many similarities with workplace monitoring as [1] defines it, and crisis management workers are facing the same problems, albeit at a different scale and for other purposes. Based on [2] and [19] I have concluded that privacy infringements exist and a too aggressive and technology-oriented evaluation may reduce performance among the audience and thereby become counter-effective.

To motivate R&E it is important that every member of the training audience gets a chance to give their informed consent, according to the definition in [17], to minimize the negative effects. It is also desirable, although not always possible; to on beforehand define exactly which questions will be involved in the evaluation and reduce data collection so that not more is collected than needed to answer those questions. By doing that, the amount of persons that have their privacy infringed reduces, as does the extent. Collected data should also be restricted to only authorized analysts.

To not violate the right of control over information about yourself, all members of the training audience should also be notified of the data that has been collected on themselves, i.e. recorded radio communication, GPS track logs, collected e-mails, etc. and be given a veto right on what can be used in further studies and who it is shared with. I believe that the mere knowledge of this right would increase acceptance among the audience and reduce anxiety as well as risk of reduced creativity. Of course the propositions herein would benefit from an empirical study, and I welcome such a study. However I warn anyone undertaking such a study to be aware that it is not always clear to the subjects what problems they are exposed to.

REFERENCES

[1] J. M. Kizza and J. Ssanyu, “Workplace surveillance”, Electronic monitoring in the workplace: controversies and solutions, J. Weckert, Eds. Hershea, PA: Idea Group Inc. Publishers, 2005.

[2] S. Miller, “Guarding the guards: the right to privacy, and workplace surveillance and monitoring in policing”, Electronic monitoring in the workplace: controversies and solutions, J. Weckert, Eds. Hershea, PA: Idea Group Inc. Publishers, 2005.

[3] D. Andersson, S. Pilemalm, and N. Hallberg, “Evaluation of crisis management operations using reconstruction and exploration”, Proc. 5th International Information systems for crisis response and management conference (ISCRAM 2008), Washington, DC, USA, May 2008, pp. 118-125.

[4] S. Pilemalm, D. Andersson, and N. Hallberg, “Reconstruction and exploration of large-scale distributed operations – multimedia tools for evaluation of emergency management responses”, Journal of Emergency Management, vol. 6, iss. 4, 2008, pp. 31-47, Weston Medical Publishing, LLC.

[5] Headquarters Department of the Army, A Leader’s Guide to After-Action Reviews (TC 25-20), Washington, DC, Sep. 1993.

[6] W. J. Rankin, F. C. Gentner, and M. J. Crissey, “After action review and debriefing methods: technique and technology”, Proc. 17th Interservice/Industry Training Systems and Education Conference (I/ITSEC 1995), Albuquerque, NM, USA, pp. 252-261.

[7] J. E. Morrison and L. L. Meliza, Foundations of the after action review process, Special Report 42. Alexandria (VA): United States Army Research Institute for the Behavioral and Social Science, 1999.

[8] Guy Carpenter & Company Ltd., “Windstorm Erwin/Gudrun – January 2005”, Specialty Practice Briefing, iss. 2, pp. 1-14, http://www.guycarp.com/portal/extranet/pdf/Speciality_Briefing_170105.pdf, retrieved 2011-05-11.

[9] T. Lay, H. Kanamori, C. J. Ammon, M. Nettles, S. N. Ward, R. C. Aster, S. L. Beck, S. L. Bilek, M. R. Brudzinski, R. Butler, H. R. DeShon, G. Ekstrom, K. Satake, and S. Sipkin, “The great Sumatra-Andaman earthquake of 26 December 2004”, Science, vol. 308, 2005, pp. 1127–1133.

[10] J. Cassuto and P. Tarnow, “The discotheque fire in Gothenburg 1998 – A tragedy among teenagers”, Burns, vol.29, iss. 5, Elsevier Science Ltd., Aug. 2003, pp. 405-416, doi:10.1016/S0305-4179(03)00074-3.

[11] Estonia Commision (Joint accident commission of Estonia, Finland and Sweden), Final report of the M/S Estonia disaster of 28 September 1994, Helsinki, Finland: Edita Ltd., 1997.

[12] S. D. Warren and L. D. Brandeis, The Right to Privacy, vol. 4, no. 5, Harvard Law Review, 1890.

[13] J. R. Aiello and K. J. Klob, “Electronic performance monitoring: A risk factor for workplace monitoring”, Organizational Risk Factors and Job Stress, S. L. Slater and L. R. Murphy, Eds. American Psychological Association, 1996, pp. 163-179.

[14] E. Palm, The ethics of workplace surveillance, Doctoral thesis in philosophy from the Royal Institute of Technology, Stockholm, Sweden, ISBN 978-91-7178-818-4, 2007.

[15] S. A. Alpert, ”Protecting medical privacy: Challenges in the age of genetic information”, Social Issues, vol. 59, no. 2, 2003, pp. 301-322.

[16] E. Merz Smith, Everything is monitored, everything is watched – Employee resistance to surveillance in Ontario call centers, MA thesis, Dept. of Sociology, Queen’s University, Canada, 2004.

[17] J. Malek, “Informed consent”, Encyclopedia of Science, Technology and Ethics, vol. 2, C. Mitcham, Eds. Detroit: Macmillan Reference USA, 2005, pp. 1016-1019.

[18] S. Clarke, “Informed consent and electronic monitoring in the workplace”, Electronic monitoring in the workplace: controversies and solutions, J. Weckert, Eds. Hershea, PA: Idea Group Inc. Publishers, 2005.

[19] B. C. Stahl, M. Prior, S. Wilford, and D. Collins, “Electronic monitoring in the workplace: If people don’t care, then what is the relevance?”, Electronic monitoring in the workplace: controversies and solutions, J. Weckert, Eds. Hershea, PA: Idea Group Inc. Publishers, 2005.

[20] W. D. Ross, The right and the good, Oxford: Clarendon Press, 1930.

 ISBN: 978-1-61208-167-0

Personal Information Systems: User Views and Information Categorization

[*] Dominique L. Scapin, [*] Pascal Marie-Dessoude, [#] Marco A. Winckler, and [*] Claudia Detraux
[(*)] INRIA, Rocquencourt, France
[(#)] IRIT, Toulouse, France
{dominique.scapin, pascal.marie-dessoude, claudia.detraux}@inria.fr, winckler@irit.fr

Abstract- This paper aims to improve Personal Information systems (PIMs) by understanding how people manage personal information items usually kept on notes, cards, agendas, etc., and on administration forms (paper or digital). The focus is both on what people say about information content, organization, trust, willingness to share and on how people categorize information. Preliminary studies (focus group and questionnaire) looked at how people describe their own use of information, and their views on future PIMs needs. They show a strong distrust towards such systems and reluctance to share personal information. Another study (card-sorting) looks at the way people assign individual information items to self-created categories. Results show a few variations in structure and naming, with a gender effect for category size. Detailed clustering and co-occurrence analyses show small differences between how people actually organize their personal information and our initial "theoretical" assignment. While the results suggest some modifications of the information structure and content, it supports the user-centric approach of the study, starting from user needs and associated documents, experimental testing and design iterations, which could be generalized for designing usable PIMs.

Keywords - personal information items, PIMs, semantic categories, naming, e-gov.

I. INTRODUCTION

Conducted within project PIMI (Personal Information Management through Internet), which goal is to develop a design environment and a deployment platform providing users with personal data access and services relevant to their needs, this study aims at gaining knowledge about the way users manage their information and services, how they see doing it in the future, and what should be the information content, structure and naming in a PIMI. Also, which items can be shared and other issues such as security and trust.

In recent years, computing technology (including Internet and mobiles) has increased capabilities for managing the large information sets needed in everyday life, professional or non-professional. Part of that information refers to personal data that users might decide to share or not through their relationships with other users (e.g., social networks) and applications (e.g., e-government services). Stimulated by the recent evolution of national governments policies towards the improvement of electronic services, e-government applications tend to require personal information for accessing public e-services such as in healthcare, taxes, housing, agriculture, education, social services [1].

Managing large sets of information is strongly related to the domain of Personal Information Management systems (PIMs), which corresponds [2] [3] to the research field addressing the way people manage their physical documents (books, notebooks, sheets, etc.), as well as their electronic documents (files, emails, Web pages, etc.), with the aim of designing tools that support the management of electronic documents (PIM tools). PIMs studies have mostly focused on very large data sets, such as the full content of user hard drive, and on search issues pointing out the large variability in people information search [3].

While the PIMs area usually covers many contexts and activities, in this paper we look at PIMs in a more specific way: the individual information items people keep on various notes, cards, forms, agendas, etc., the ones that are personally attached to us, that we use in our every daily life, both professional and non-professional (for administrative, social, leisure purposes, etc.). It is not the full set of available files. Besides, we look at the intuitive way people organize their personal information, with or without computer systems. Concerning personal information bits currently scattered many places, there is little research with a user-centric approach, with the view that users-based knowledge might help specifying computer-based tools.

This paper starts with a review of literature and publicly available tools. Then, preliminary studies briefly report on documents analyses and on what people say and wish about their personal information. A section describes the method, tool, procedure and participants of a card-sorting study on people intuitive organization of personal information. The results and their impact towards a future PIMI structure are presented. The conclusion summarizes the study, identifies its limits and provides insight on future research work.

II. RELATED WORK

The context is e-gov. (short for electronic government), a diffused neologism used to refer to the use of information and communication technology to provide and improve government services, transactions and interactions with citizens, businesses, and other arms of government [4].

In the area of e-gov., a literature review [5] showed little specifically on human factors in HCI (Human-Computer Interaction). Studies identified dealt mainly with user needs and accessibility (a major topic in e-gov. HCI, including studies on older people), the applicability of HCI results to e-gov., ad hoc interaction novelties (e.g., animated faces), ad hoc methods (e.g., on document exchange and scenario

 ISBN: 978-1-61208-167-0

planning), issues of user involvement and requirements, user acceptance, and patterns.

About PIMs, a more substantial body of knowledge in many different settings has begun to pile up in recent years [6] [7]. There is a lot of technical aspects such as: data synchronization across devices, version control [8], file management and applications [9], collective work and file sharing [10], novel user interface paradigms and mobility [11] [12], ontologies [13], and tools based on information association [14] [15].

There are also usability, citizen-centric studies, mainly about: privacy and security [16], hierarchical files structure issues and proposal for a tagging mechanism [17], studies following-up on [3, op. cit.], such as: further empirical investigation on ways to improve information searching [18], investigation of the role of personal notes [19], contextual use of PIMs [20], tool evaluation [21], and call for more user-centered studies, long-term studies on the evolution of user information practice from one work context to another, from a role to another [22]. In light of our goals, a few points can be selected from these studies.

One point is that hierarchical structure is the most used and preferred by users [23] [24] [25] [26]. The latter study actually shows that users built an ownership and control feeling about their data, probably due to long use of such tools. However, users have difficulties in creating consistent structures and naming items categories. More specifically, categorizing and naming new items in an existing structure seems to be difficult and represent a high cognitive load [24, op. cit.], [25, op. cit.]. Adding contextual data, such as tags, may help, but tagging may vary from one user to the other, and will not solve consistency issues, particularly when information spaces are to be shared, even partly. An interesting addition to contextual data from sensors (GPS, GSM, and movement) has shown to help find images within a collection [27].

About search strategies, users tend to first explore the structure, and use search tools only afterwards. Even though it may be explained by lack of user knowledge [26, op. cit.], the lack of flexibility of these tools may still apply [3, op. cit.].

One particularly pervasive PIM problem is information fragmentation [28], i.e., when information related to a single task is scattered across several different applications and environments. A typical example is project information, where specifications may be in a Word document, budget in an Excel file, communication with the customer and the project manager may be in emails, and other resources may reside in Intranet or the Web. A project member may have seen all these documents but may later have trouble re-accessing them. In [28, op. cit.] it is argued that grouping related information is a central PIM activity currently hindered by the artificial separation imposed by the different applications. In addition, history and versioning must be dealt with.

On the practical side, [28, op. cit.] suggest 5 factors that may hinder PIMs use: visibility, integration, co-adoption, scalability, return on investment. In absence of visibility, when the PIM is not always visible to the user, the tendency is to forget it, as well as the data already stored. Integration: when not integrated with the other tools, it can be underused. Co-adoption: for user cooperation, share and synchronization of data (e.g., appointments, agenda) is required, if not available the PIM might not be used. Scalability: the PIM tool must allow scaling (e.g., more data, projects). Return on investment: if the tool requires a large learning effort, it will not be used. Guidelines are also offered by [26, op. cit.], along three types of strategies: piling, filing, and structuring.

Concerning information transfer, results of a survey with 47 participants [30] showed that the main forms of PIM storage are computers, then external disk drives. For Web sites, the ordered storage preferences are bookmarks, email, paper. When data transfer, it is done mainly with email and memory sticks, and in some cases on web sites. The main difficulty is finding the files. The results also point out the important role of email systems in storage, sharing, search, and file exchange between computers and other devices.

In our research, the attempt is to complement these earlier findings, first on what people say (in terms of information items, of current practice, of shareability, etc.), but also focus on a novel issue: what people do intuitively with an existing set of unstructured information items.

We also looked at 15 tools [31] that claim to support personal information management. Most tools (Tools # 1, 2, 3, 4, 5, 7, 8, 11, 12, 13, 15) offer an agenda, a calendar, a contact list, a keyword based search tool, a centralized password management function, and notes editing. In addition, a few tools offer more sophisticated functions such as: a sort of mind-mapping allowing to represent user's thoughts (Tool # 2), a text-based card information management system (Tool # 6), a system for managing archives (Tool # 9), a document and pictures storage system dedicated to Android-based mobile phones (Tool # 14), and a note management system that coordinates notes (containing files) between (Mac, PC, mobiles) platforms (Tool # 10). Few of these tools are available on line (e.g., Tools # 7, 9, 15). Most others run on personal computers, except for Tool # 14 that runs only on mobiles and Tool # 10 that runs on personal computer, internet and mobile. Synchronization is provided by Tools # 10 and 15.

Most user interfaces are rather similar. Some allow tailoring, mainly on colors, window size, language, menu position. A few points concern usability issues. For instance, concepts may be difficult to grasp due to naming (e.g., Tools # 1, 4, 7). In general, no predefined structures, or default patterns are included, or are unusable for new user entries, which might make it difficult for a large public. Also, the coverage of tools, whenever items and categories are proposed, does not include many of the citizens' information such as identities, health, finance, work, etc. Some of the associated information search tools are powerful (e.g., Tools # 1, 2, 7), but a bit cumbersome (many forms and choices). Some tools offer also tags, labels, particularly for notes, contacts and events (e.g., Tools # 1, 2, 3, 6, 10, 13, 15).

Overall, useful insight can be extracted from the references in this related work section. However, few studies concern the information content of PIMs, and very few include a full user-centric process approach, such as

 ISBN: 978-1-61208-167-0

proposed in current standards, particularly ergonomics standards.

III. PRELIMINARY STUDIES: WHAT DO PEOPLE SAY ABOUT PERSONAL INFORMATION

An initial step has been the analysis of 9 administrative forms relating to international actions, students grants, solar heating incentives, livestock diseases, sport incentives, etc. available from French government and administration.

The analysis showed a very large diversity of information items (500 + 200 synonyms). 21 different topics were identified. This information content was classified and later supported the design of questionnaires and card-sorting material.

A. Focus groups

With three focus groups (a group with 6 participants from a large industrial group, and two 3rd. year university students groups, with respectively 7 and 9 participants), a study investigated the issue of electronic information storage for personal information, as well as the issue of shareability.

Besides providing a number of candidate information items for a personal information space, the main issues covered concerned: a strong distrust of electronic storage of personal information (hacking, piracy), the time needed to proceed to a single electronic storage of personal information, and their reluctance to share personal information.

B. Online survey

A questionnaire survey investigated current practice along the same issues: electronic personal information storage, shareability, etc, with the addition of people's view on their future personal information space. A self-administered questionnaire was available on Internet [32] using the tool SurveyMonkey [33]. The survey was voluntarily focused on personal information used regularly by the public.

The survey was answered by 30 participants: 14 clerks and 16 3rd. year university students. 29 respondents were women.

First, users were asked about their profile and how they dealt currently with their personal information.

Secondly, they were presented with a set of 114 personal information items organized into 9 categories and 26 sub-categories. Figure 1 shows the "theoretical" category structure, a pre-defined structure based on preliminary studies. Categories are organized in a horizontal menu whilst items within a category in the vertical menus options below.

For each sub-category users were asked to identify information items that should belong, provide alternatives names if unsatisfactory, express willingness to share that information with others, and tell if each sub-category should be part of a personal information space. Thirdly, users browsed all items, and were asked to point any category that could be missing and to comment on PIMs advantages/drawbacks, potential uses, willingness to share their PIMS with administrations, views on tools for automatic filling electronic forms.

MY PERSONAL INFORMATION SPACE								
My Identification	My Family	My Health	My Professional Activities	My Transportation means	My Finances	My Logins & Passwords	My Agenda	My Contacts
My Identity	My Family Status	My Health Coverage	My Career	My Personal transportation	My Bank	My IDs, Logins and Passwords	My Personal Agenda	My Personal contacts
My Contact Information	My Parents	My Physician	My Current Job	My Public Transportation	My Income and Social Benefits		My Professional Agenda	My Professional Contacts
My Identity documents	My Spouse / Partner	My Medical Records			My Investments			
My Biometric Data	My Children				My Loans			
					My Wealth			
					My Income Taxes			

Figure 1. Structure of the "theoretical" personal information space

Due to space, results on current and future use are not reported here, except to mention a strong distrust for such systems and reluctance to share information. Focusing on the categorization and naming issues, the results support the study material. Regarding information items categories, the participants tend to agree (79,2%) with the categories offered. For the remaining 20,8% the few categories suggested are: Hobbies", "Insurance", "Music", "Sports", "Union activities", as well as "Spending". Regarding naming, the participants did not make much suggestion. All items are well accepted, except 5 (out of 114) for which proposals are made.

IV. HOW DO PEOPLE ORGANIZE THEIR PERSONAL INFORMATION ITEMS

Another study focused on how participants organize information items. They were asked to create their own "boxes", i.e., information categories in which to insert their personal information items, using a Card-Sorting technique.

A. Material, procedure, participants

Card-Sorting is a way of gaining insight on categorization and mental models about information architecture that can be described by means of small cards [34]. It starts with writing each statement on the information architecture on a small card. Then, participants are asked to sort a set of cards with words or pictures into piles of similar cards. Participants may be asked to provide labels for the card piles they have created or they may be provided with pre-defined labels and asked to match the cards to them.

The Card Sorting tool: currently several card-sorting tools are available for enabling users to classify items on a computer instead of using paper cards. Most tools run under Web platforms, allowing card-sorting studies to be administered remotely. The results provided by online tools are quite similar to studies run using paper-based cards [35]. For this study, we used the tool WebSort [36]. The tool was initialized with our 114 personal information items (in alphabetical order, random assignment not being manageable with the tool). Users moved items from the list presented at the left side to the right area, creating groups of items they named as categories.

Material: 114 personal information items focusing on everyday life aspects involving health, banking, social welfare, citizenship administrative papers, etc.

 ISBN: 978-1-61208-167-0

Procedure: a call for participation was distributed through web sites and email lists. The study material was accessible from an Internet address. After a short definition of "Personal Information", of "Personal Information Space", and a description of the study goals, and tool functions (creating boxes, naming, renaming, etc.), participants were instructed to run their individual session at their own pace. Once the session was completed and saved, the participants were asked to fill a questionnaire (participants' profile, current ways of managing your personal information, views/ suggestions on a future personal information space).

Participants were recruited through various professional social networks and email lists. Aside 6 initial answers used for pre-testing, and 3 beta-testing answers, 56 participants responded to the card-sorting study. Due to incomplete answers (from 5 to 112 unsorted items, or no answer to the questionnaire), 13 participants were excluded from data analysis.

The characteristics of the remaining 43 are: 32 male participants (74%) and 11 female (26%), 33 in the 18-39 y. age bracket (77%), 9 in the 40-59 y. (21 %) and 1 in the 60-74 y. (2%), 9 are single (21 %) while the 34 others are not (78%) [18 married and 16 shared living], 14 participants have children (33%), while 29 do not (67%), 31 participants are employees (72%), 5 self-employed (12%), 4 students (9%), 2 unemployed (5%) and 1 retired (2%), 41 from France (95%), 1 USA and 1 Belgium.

B. Card sorting results

The average time for a full session (card-sorting and questionnaire) was 42 minutes: for the Card-Sorting part only, average duration was 35 minutes.

The next sections describe the categories (from now on called "boxes") created by the participants, their size and content, variations in naming, clustering aspects, and the analysis of the role of the participants' characteristics.

Boxes created: overall, 43 participants each classified 114 information items, for a total of 4802 items. Together, 500 boxes were created. Each participant created from 5 to 22 boxes (mean= 11.627, sd= 4.434, median= 11.000). In each box, they included 1 to 53 items per box (mean= 9.739, sd= 7.291), i.e., 34 different sizes.

Four sizes have been distinguished: “Very Large Size Boxes” > 21 items (and < 1 to 6 > boxes that size), “Large Size Boxes” < 9 to 21 items > (and < 6 to 28 > boxes that size), “Midsize Boxes” < 3 to 8 items > (and < 34 to 50 > boxes that size - *most numerous*), and “Small Size Boxes” < 3 items. Most topics are distributed across boxes, big, mid-size, and small, without much recognizable patterns.

These variations are further illustrated Figure 2 showing both (ordinate) size of boxes (i.e., number of items in boxes) and (abscissa) number of boxes each size (i.e., equal number of items in a box).

One can see that there is a rather regular parallel increase/decrease for the boxes size/number of items, respectively, until 25 items per box, while from 27 items, the curve gets more erratic for the number of large size boxes.

The next section looks more closely at the boxes content.

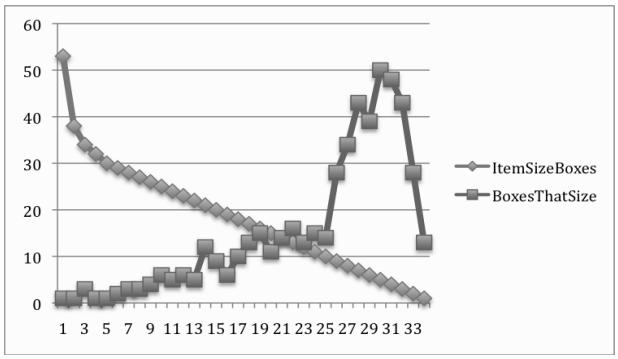


Figure 2. Curve # items per box X # boxes of each size

Lexical variations: naming of boxes will be later analyzed, particularly to help define naming of categories in the future PIMI tool. We simply mention here a few elements of variation.

- Imprecise naming, e.g., *later, confidential, others.*
- Typographic variations: caps/ lowercase, capitalized, with/ without accents, typos.
- With possessive or not, e.g., *(My)health, (My)family.*
- Singular vs. plural, e.g., *finance(s), address book (s).*
- Syntax: verb or noun or adjective, e.g., *administrative, administration.*
- Synonyms and abbreviations versus names.
- A quite large area of variation concerns the use of several terms together covering different aspects.

Conceptual variations: besides the fact that smallest size boxes cover more specific concepts than the larger ones, we looked at their scope and meaning. We distinguished Single concepts, Multiple Concepts (Explicitly Grouped by Name), and Higher-Level Concepts. The way boxes group items is quite variable in terms of concept scope, which confirms the analysis on boxes sizes.

These linguistic and conceptual variations have given some hints on default names for a future editable PIM, but also point at conceptual associations that are discussed in the clustering analyses. For instance, the frequency of identical "concept names" is quite high for boxes such as "health", "bank", "professional life".

Clustering analysis: a multiple correspondence analysis (MCA), using the SPAD tool [37] was performed with 9 classes (as in theoretical structure). Multiple Correspondence Analysis (MCA) is a method for nominal or categorical data sets (e.g., [38] [39]). The goal is to visualize a data set by representing data as points in a low-dimensional Euclidean space. This procedure is similar to principal component analysis for categorical data. MCA is also an extension of simple correspondence analysis (CA) in that it is applicable to a large set of categorical variables. In this case MCA is a CA on the Burt table formed from these variables. As in factor analysis methods, the first axis is the most important dimension, the second axis the second most important, and so on. The number of axes to be retained for analysis is determined by calculating the eigenvalues (a set of scalars associated with a linear system of equations, i.e., a matrix equation, sometimes also known as characteristic roots or

 ISBN: 978-1-61208-167-0

values, proper values, or latent roots). In our application we have retained 20 axes. On this low-dimensional Euclidian space we applied a hierarchical method. Hierarchical algorithms find successive clusters using previously clusters. This algorithm chosen is an agglomerative ("bottom-up") algorithm. This algorithm begins with each element as a separate cluster and merges them into successively larger clusters. An important step in most clustering is to select a distance measure, which will determine how the similarity of two elements is calculated. We have selected the Euclidian distance on the 20 factors selected by the MCA and the aggregative criterion used is the Ward's criterion [40].

Using such hierarchy in the Ward sense, 9 classes were detected. It shows overall a rather consistent category assignment to items, but some differences with the "theoretical assignment". The initial clustering (see Table I) started by leaving out 34 individual items, while establishing 33 initial groups (called level 1 clusters). They vary in item numbers content: 24 groups with 2 items, 7 with 3 items, and only 1 with 5 or 6 items.

TABLE I. NUMBER OF ITEMS IN INITIAL CLUSTERS

# Items	**1**	**2**	**3**	**4**	**5**	**6**	**Total**	
# Groupings		24	7	0	1	1	**34** single	**33** grouped
# Total Items	**34**	48	21	0	5	6	**114** items (80 grouped)	

Distribution of the number of items per level is shown Figure 3. Clustering varies little up to level 5 in terms of items numbers. There is a drop for levels 6 to 9 that actually tend to group item clusters from lower levels.

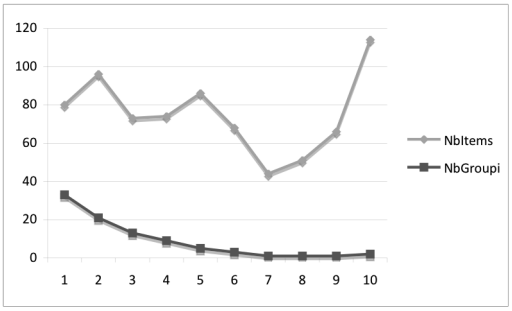


Figure 3. Curve items per level

To detail results, the 9 classes are seen at the incremental level, from "leaves" to "root" in the SPAD dendrogram. The clustering level refers to the rank at which the information items are grouped together in the clustering analysis: each time individual items or clusters are grouped, it adds a level (incrementing by 1 the previous highest level). Otherwise said, the more levels within a class, the more variations in boxes assignment. It is a hint on coherence of certain groups compared to others (e.g., many levels within the finance area, while few levels in the taxes, or health areas).

Looking more closely at the tree, an illustration is provided Figure 4 for non-health clusters and Figure 5 for health clusters. All grouped items are "boxed", individual items are not. All nodes show a level (L1 to 10). The lowest level is L1 (minimum items is 2, listed with "+" to associate them). All level 1 nodes show in addition a parenthesis referring to the level rank as used in the explanatory text below. The 9 SPAD classes are identified with "★ #" (8 non-health, 1 health). The main observations are the following.

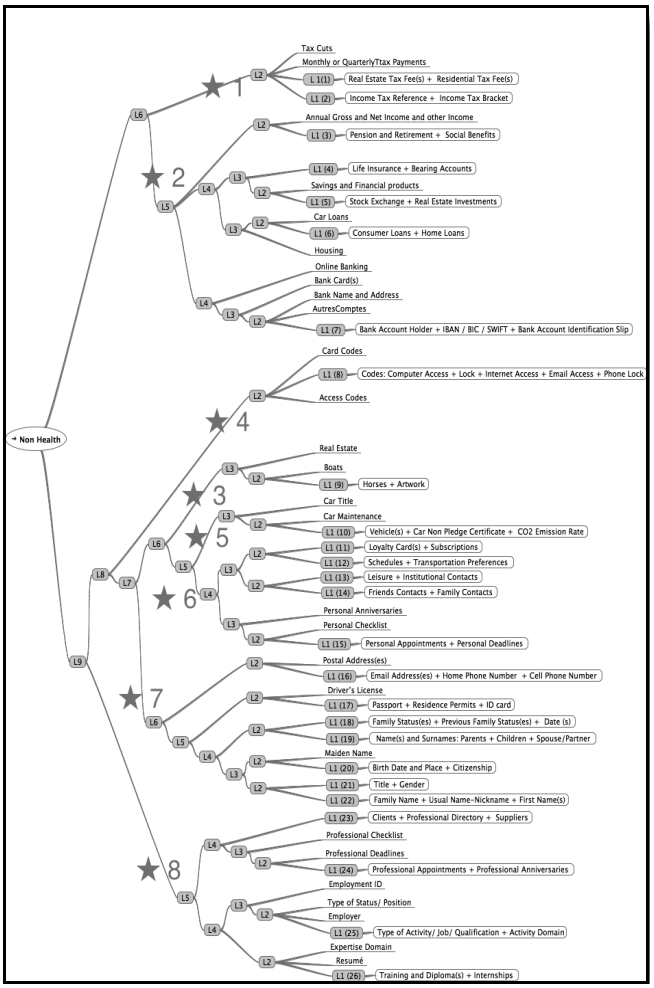


Figure 4. Non-Health Clustering Tree

A **first** class of 6 items corresponds to taxes. It is a quite consistent group, not been found together with other financial considerations. This suggests it is a specific concept for the participants.

A **second** class of 19 items deals with other financial aspects: bank (7 items), income (3 items), investments (4 items), loans (4 items), and savings (1 item). This actually contradicts the theoretical assignment by excluding patrimonial aspects.

A **third** class of 4 items relate to a "patrimonial" category, distinct from other financial aspects. Unlike theoretically, monetary savings do not belong, nor vehicles meant as collectible, which may not have been clear.

A **fourth** class of 7 items groups all codes. This cluster is distinct from the other groupings: it is only grouped with other clusters at a much higher level.

 ISBN: 978-1-61208-167-0

A **fifth** class of 5 items groups information related to car information (4 items) to which a *Vehicle* item was added (meant to be vintage/ collectible cars, but not clearly expressed in card sorting explanations). This corresponds to the theoretical "My Personal Transports" with the addition of *Vehicle*, but without *Driver's License*. This is a case of potential redundancy within a PIM, a driver's license being at the same time a vehicle, and a proof of identity.

A **sixth** class of 12 items groups a more heterogeneous set dealing with: public transportation (4 items), personal dates and appointments (4 items), and leisure (3 items), adding *Institutional contacts*, probably viewed as the ones needed in every day life. Overall, this cluster shows a large number of different clustering levels, which makes it a bit unstable compared to others. Again, it may mean that some information items should be made redundant, depending on their context of use.

A **seventh** class of 22 items groups various aspects of people identification: personal identity (8 items), contact information (4 items), identity documents (3 items), family status (3 items), information on relatives (3 items), and *Driver's License,* which seems to be considered as part of a person's identity. Also, it does not include *Health records,* which was in the theoretical assignment, but meant for children health.

An **eight** class of 16 items groups all professional information: career (4 items), job (5 items), but also professional agenda (4 items), professional/clients contacts (3 items). Participants tended to group all professional items, rather than distinguishing documents/ status items and contacts aspects.

Finally, a **ninth** class of 23 items joins all health aspects, whether biometric data (5 items), or health analyses (12 items), or administrative and health contacts (5 items), and children *Health Record*.

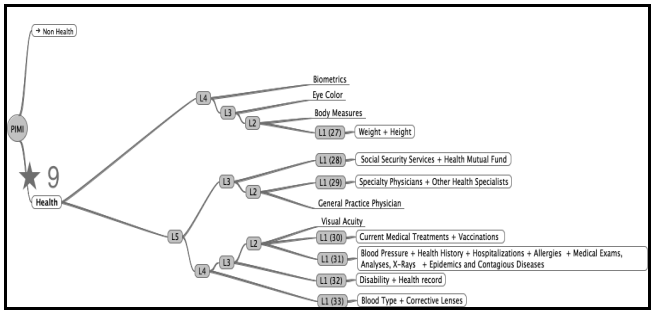


Figure 5. Health Clustering Tree

Co-occurrence of items in boxes: the analysis searched all cases ("frequent sets") for which at least 80% of the participants (i.e., at least 35) put a set of items into the same box. This resulted in 1041 sub-sets of items. To reduce that large collection, we looked at "closed frequent sets" within these "frequent sets". A "closed frequent set" is a set for which all its sub-sets are frequent, but which is not itself contained in another "frequent set". By topics, the items were grouped from largest to smallest number of participants within largest to smallest sets within a topic. The maximum of items found together is 9, minimum being 2. The analysis is consistent with previous clustering analyses. It also shows that the primary topics under which information items are being put together on a regular basis (highest co-occurrence topics, ranked first by number of participants, and within, by number of items) relate to health (13), bank/ finance (12), and identity (12), followed by work (7) and codes (6), the last ones being taxes (2), loans (1), telephone (1), agenda (1).

Differences due to participants' characteristics: the focus is here, per participant, on the number of boxes, the size of the smallest boxes, the average size of the boxes, as well as their maximum size (see Table II).

TABLE II. FISHER TEST AND ANOVA ON GENDER

VARIANCE ANALYSIS :				(Fisher) Student's T		ANOVA				
	Value	Average	Std-dev	Test		Variance decomposition		Significance level		
Nb_Boxes	W	13,8182	**5,3632**	Fisher **T**	1,9073	Source	Sum of square	Statistics	Value	Proba
	M	10,8750	**3,8834**	**df**	10/31	BSS	70,9101			
	All	11,6279	4,4348	**p-value**	0,1649	WSS	755,1364	Fisher's **F**	3,850054	**0,056556**
						TSS	826,0465			
Min_ite ms	W	**2,6364**	1,6293	Fisher **T**	5,8235	**Source**	**Sum of square**	**Statistics**	**Value**	**Proba**
	M	**4,3438**	3,9318	**df**	31/10	BSS	23,8637			
	All	3,9070	3,5511	**p-value**	**0,0056**	WSS	505,7642	Fisher's **F**	1,934522	0,171768
						TSS	529,6279			
Avg_ite ms	W	9,5776	3,8759	Fisher **T**	1,4673	**Source**	**Sum of square**	**Statistics**	**Value**	**Proba**
	M	11,9877	4,6950	**df**	31/10	BSS	47,5487			
	All	11,3712	4,5802	**p-value**	0,5339	WSS	833,5487	Fisher's **F**	2,338791	0,133867
						TSS	881,0973			
Max_i	W	23,3636	7,7236	Fisher **T**	1,0351	**Source**	**Sum of square**	**Statistics**	**Value**	**Proba**
	M	23,1563	7,8581	**df**	31/10	BSS	0,3521			
	All	23,2093	7,7323	**p-value**	1,0180	WSS	2510,7642	Fisher's **F**	0,005749	0,939928
						TSS	2511,1163			

Out of the participants' characteristics (gender, age, marital status, children, activity, location), only gender showed a significant role. It mainly means that women create more boxes than men, with a higher variability (Anova F, $p < 0.0565$, women std. dev. 5.3632, men std. dev. 3.8834). However, women tend to create less small boxes (Student T, $p < 0.0056$), women average 2.6364, while men average 4.3438. Otherwise said, women tend to use more boxes and to vary more in their size and content.

C. *New information category structure*

Results showed that our items list, initial classification, and labels were quite satisfactory for participants.

However, the analyses showed some discrepancies suggesting a new candidate structure along: Identity & Contacts (personal Identity, Identity papers, personal Contacts), Work (current Work, Career, professional Contacts), Agenda, Contacts & Transports (personal Agenda, Contacts, individual Transports, public Transports), Codes & Passwords, Finances (Income and social benefits, Investments, Loans, Bank accounts), Taxes, and Health (Social security & Mutual Funds, physicians, and medical Records).

Compared to the previous structure, there are 7 categories instead of 9. "My Family" is removed. "My Agenda", "My Contacts", and "My Transportation" are grouped, and a specific category "Taxes" is created. Agenda may be questionnable as distinguishing personal versus professional may not make sense if a future tool includes a common calendar. In addition, categories and sub-categories were reordered according to expected frequency and initial tool set-up order. Information items were reduced based on assignment stability, but more pragmatically on expected usefulness (in a new experiment, young students, may not need in their PIMI tool topics such as patrimony or children health). In terms of naming, the use of the possessive will not be pursued and a number of category and information names

 ISBN: 978-1-61208-167-0

will be changed to reflect the most frequent name generation by the study participants. Flexibility is another issue of prime importance. Despite any efforts at designing a good structure, good naming, there is probably no single solution: a future PIMI tool probably should offer a user tailorable structure and content, as well as flexible search features.

D. Questionnaire results

The participants input to the questionnaire after the card-sorting exercise is quite rich, and will need further qualitative analyses. The main quantitative results, rather consistent with the previous survey, are as follows:

- 86.5% participants use both paper and electronic storage, while 7.7% only electronic and 5.8% only paper. 91.7% use a personal computer, 27.1% a cellular phone, 10.4% an iPhone, 8.3% a smartphone, 8.3% a PDA, and 18.8% other means.
- Only 30.6% use some information protection method for their personal information while 69.4% do not.
- 46.9% currently share some of their personal information, while 53.1% do not. They share it with spouse (62.5%), administration (37.5%), social networks (34.4%), family (31.3%), friends (31.3%), employer (21.9%), colleagues (18.8%), and 9.4% others. The topics for which they show reluctance to share are codes, health-related information, and income.
- 70.6% organize their personal information into categories, while 29.4% do not. When stated, the categories correspond either to the ones proposed or to their naming of the card-sorting boxes, with the addition of a few categories such as “bills”, "bicycle", "religion", "politics".
- For a future information space, 65.2% did not offer new categories, while 34.8% added a few new ones, e.g., friends birthday, food recipes, computer IP, bills, music, photos and videos.
- 71.4% see both advantages and drawbacks in a future information space, while 22.8% see only advantages, and 5.7% only drawbacks. Those mainly concern trust and security, while advantages concern centralization and ease of access. Regarding expected functionalities, the top ones are searching, storage, filtering, and accessibility.
- 72.3% are worried about sharing their information with administrations, while 27.7% are not. The main concerns, relate again to trust and security.
- 55% are in favor of automatic form-filling, while 42.5% have mixed feelings, and 2.5% are against it. The main benefit mentioned is time saved, while the concern relates to strict selection of information to be auto-filled (security and confidentiality issues).

V. CONCLUSION

This paper reported novel user-centric work in the area of PIMS and e.gov., focusing on the personal information bits currently scattered many places, electronic or paper, rather that on disk file management issues. The methods used attempted to identified what people say and what people do about their personal information. These methods have limits related to remote experiments through Internet, to the card-sorting tool (management of redundancies, sub-categories, monitoring participants’ modification strategies).

However, the results support the user-centric approach of the study, starting from user needs and associated documents, experimental testing and design iterations, which could be generalized for designing usable PIMs. On the practical side, the results also suggest a few modifications of the information structure and content to be further tested in the PIMI project.

The design of tools for supporting citizen to use and share personal information is a complex task. Some of the issues are trust and willingness to share information which is of prime importance to users (even though many of them have lesser concerns when posting very private material on Facebook or other social networks), coping with users behavior variability and preferences, setting up proper procedures for information exchange in e-gov. contexts, reducing the information fragmentation issue, such as making sure different tools and environments allow consistent and synchronized use resources, and providing efficient search tools, with queries adapted to the users.

In our research, the next step is to test a mockup system based on the previous findings. The mockup will be tested in depth, with users being monitored, with tasks to perform, with usability measurements, both objective and subjective. One underlying idea is to explore how people can actually tailor their own information space, in terms of structure and naming, but also in terms of sharing parameters, and associated (flexible and contextualized) search tools.

Later on, the current static view (i.e., about content, information items and categories) will be extended to a dynamic view that will include procedures (i.e., using personal information items to manage one's personal space and local information transfers, as well as to fill manually or automatically e-gov. forms). This will concern service composition, building ontologies, associating information items related documents (e.g., file copy of driver's license, of administrative documents).

An additional goal will be to formalize the design process for delivering validated new e-gov. content: context of use study (e.g., documents), user needs gathering (ideas through focus groups, facts and opinions through questionnaires), concepts definition (design step), subjective testing (through questionnaires), content assessment (through card-sorting and questionnaire), prototype design and user testing.

Hopefully, in the end, this will contribute to more citizen-centric personal information systems.

VI. ACKNOWLEDGMENTS:

Thanks to Y. Lechevallier for his expert contributions and advice on statistical analyses, as well as for the numerous stimulating discussions. Preliminary studies have been funded by the ANR project "MyCitizSpace". The card-sorting study has been funded by the ANR project "PIMI".

 ISBN: 978-1-61208-167-0

VII. REFERENCES

[1] Calvary, G., Serna, A., Coutaz, J., Scapin, D. L., Pontico, F., and Winckler, M.. "Envisioning advanced user interfaces for e-government applications: a case study". In Practical Studies in E-Government: Best Practices from Around the World. Assar, S., 2010, Chapter 12, pp. 205-228.

[2] Bergman, O., Boardman, R., Gwizdka, and J., Jones, W. "Personal information management". SIG at CHI 2004, April 24–29, 2004, Vienna, Austria, pp. 1598-1599.

[3] Blanc-Brude, T. and Scapin, D. L. "What do people recall about their documents? Implications for desktop search tools", Proc. IUI 2007, Honololu, HI, USA, January 28-31 2007. pp. 102-111.

[4] http://en.wikipedia.org/wiki/E-Government

[5] Scapin, D. L. "E-government HCI: a genuine research field?" Proc. DEGAS'2009, In conjunction with INTERACT'2009, August 24, 2009, Uppsala, Sweden, pp. 33-37.

[6] Teevan, J., Jones, W., and Bederson, B.. "Special issue: personal information management". Communications of ACM 49, 1, January 2006, pp.40-43.

[7] Jones, W.and Teevan, J. "Personal information management", 2007, Seattle, WA: University of Washington Press. Jones, W., Teevan, J. (Editors), ISBN 978-0-295-98737-8.

[8] Tungare, M., Pyla, P. S., Pérez-Quiñones M,. and Harrison, S. "Personal information ecosystems and implications for design", Technical Report cs/0612081, ACM Computing Research Repository, 2006.

[9] Sauermann, L., Grimnes, G. A., and Roth-Berghofer, T. "The semantic desktop as a foundation for PIM research". PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[10] Collins, A. and Kay, J. "Collaborative personal information management with shared, interactive tabletops" . PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[11] Jetter, H. C., König, W., Gerken, J., and Reiterer, H. "ZOIL a cross-platform user interface paradigm for personal information management". PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[12] Woerndl, W. and Woehrl, M. "SeMoDesk: towards a mobile semantic desktop". PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[13] Katifori, A., Vassilakis, C., Daradimos, I., Lepouras, G., Ioannidis, Y., Dix, A., et. al. "Personal ontology creation and visualization for a personal interaction management system". PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[14] Chau, D. H., Myers, B., and Faulring, A. "Feldspar: a system for finding information by association". PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[15] Diehl, J. "Associative personal information management". CHI 2009, New York, NY, USA, pp. 3101-3104.

[16] Karat, C-M., Brodie, C., and Karat, J. "Usable privacy and security for personal information management". Communication ACM 49, 1 January 2006, pp. 56-57.

[17] Hsieh, J. L., Chen, C. H., Lin, I. W., and Sun, C. T. "A web-based tagging tool for organizing personal documents on PCs". PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[18] Fuller, M., Kelly, L., and Jones, G. J. F. "Applying contextual memory cues for retrieval from personal information archives". PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[19] Bernstein, M., Van Kleek, M., Karger, D., and Schraefel, M. C. "Information scraps: How and why information eludes our personal information management tools". ACM Trans. Inf. Syst. 26, 4, 2007, pp.1-46.

[20] Bergman, O., Beyth-Marom, R., and Nachmias, R. "The user-subjective approach to personal information management systems design: evidence and implementations". J. Am. Soc. Inf. Sci. Technol. 59, 2 (January 2008), pp. 235-246.

[21] Gonçalves, D. and Jorge, J. A. Now, "It's personal! evaluating PIM retrieval tools". PIM workshop CHI 2008, April 5-6, 2008, Florence, Italy.

[22] Tungare, M. "Understanding the evolution of users' personal information management practices". INTERACT'07, Vol. Part II. Springer-Verlag, Berlin, Heidelberg, pp. 586-591.

[23] Henderson, S. and Srinivasan, A. "An empirical analysis of personal digital document structures", Proc. HCI International 2009, Berlin, Heidelberg: Springer-Verlag, 2009, pp. 394–403

[24] Evequoz, F. and Lalanne, D. "I thought you would show me how to do it - studying and supporting PIM strategy shanges". PIM Workshop, ASIS&T 2009, Vancouver, BC Canada, Nov. 5 - 7 2009, pp. 35-42.

[25] Voit, K., Andrews, K., and Slany, W. "Why personal information management (PIM) technologies are not widespread". PIM Workshop, ASIS&T 2009, Vancouver, BC Canada, Nov. 5 - 7 2009, pp. 60-64.

[26] Henderson, S. "Guidelines for the design of personal document management user interfaces". PIM Workshop, ASIS&T 2009, Vancouver, BC Canada, Nov. 5 - 7 2009, pp. 65-72.

[27] Jakob, L. and Maciej, L. "Using mobile phone contextual information to facilitate managing image collections". PIM Workshop, ASIS&T 2009, Vancouver, BC Canada, Nov. 5 - 7 2009, pp. 73-75.

[28] Karger, D. R. "Unify everything: it's all the same to me", in Personal Information Management, (Jones & Teevan, eds.) University of Washington Press, Seattle, WA, pp. 127-152.

[29] Jones, E., Bruce, H., Klasnja, P., and Jones, W. "I give up! Five factors that contribute to the abandonment of information management strategies". ASIS&T 2008, 4,1, pp. 1-6.

[30] Capra, R. A "Survey of personal information practices". PIM Workshop, ASIS&T 2009, Vancouver, BC Canada, Nov. 5 - 7 2009, pp. 2-5.

[31] 15 Tools for personal information management: #**1** www.efficientpim.com, #**2** www.thebrain.com, #**3** www.essentialpim.com, #**4** www.winpim.com, #**5** www.lifemanagerpro.com, #**6** www.azzcardfile.com, #**7** www.pimonline.com, #**8** www.pimone.com/pimone.htm, #**9** www.myarchivebox.com, #**10** www.evernote.com, #**11** http://code.google.com/p/keynote-nf, #**12** http://www.treepad.com, #**13** www.milenix.com, #**14** www.android-software.fr/pocket-docs, #**15** www.gmail.com

[32] http://ihcs.irit.fr/winckler/MyPIM/index.html

[33] http://www.surveymonkey.com

[34] Fincher, S. and Tenenberg, J. "Making sense of card sorting data". Expert Systems, 22(3), 2005, pp. 89 – 93.

[35] Bussolon, S, Russi, B., and del Missier, F. "Online card sorting: as good as the paper version". ECCE 2006, Sept 20-22, Zurich, Switzerland, pp. 113-114.

[36] http://websort.net

[37] Morineau A. et Morin S. " Pratique du traitement des enquÍtes. Exemple d'utilisation du logiciel SPAD". Paris, 2006. (ISBN 2-9525948-0-5)

[38] Benzécri, J.-P. "L'analyse des données. Volume II. L'analyse des correspondances". Paris, France: Dunod, 1973.

[39] Greenacre, M. and Blasius, J. (editors) (2006). "Multiple correspondence analysis and related methods". London: Chapman & Hall/CRC.

[40] Ward J. H., "Hierarchical grouping to optimize an objective function", J. American Statistical Ass., n°64, 1963, pp. 236-244.

 ISBN: 978-1-61208-167-0

Diversity in Recommender Systems

Bridging the gap between users and systems

Laurent Candillier[*], Max Chevalier[**], Damien Dudognon[*,**], Josiane Mothe[**,***]

[*]OverBlog
WikioGroup
Toulouse, France
firstname.name@wikiogroup.com

[**]IRIT, UMR5505
University of Toulouse
Toulouse, France
firstname.name@irit.fr

[***]IUFM
University of Toulouse
Toulouse, France
firstname.name@toulouse.iufm.fr

***Abstract*—Recommender systems aim at automatically providing objects related to user's interests. The angular stone of such systems is a way to identify documents to be recommended. Indeed, the quality of these systems depends on the accuracy of its recommendation selection method. Thus, the selection method should be carefully chosen in order to improve end-user satisfaction. In this paper, we first compare two sets of approaches from the literature to underline that their results are significantly different. We also provide the conclusion of a survey done by thirty four students showing that diversity is considered as important in recommendation lists. Finally, we show that combining existing recommendation selection methods is a good mean to obtain diversity in recommendation lists.**

Keywords-Information Retrieval; Recommender System; Document similarity; Diversity

I. INTRODUCTION

Information Retrieval (IR) usually sorts the retrieved documents according to their similarity with the user's query. Doing so, they assume that document relevance can be calculated independently from other documents [10]. As opposed to this assumption, various studies consider that a user may prefer to get documents that treat of various aspects of her information need rather than possibly redundant aspects within documents [6][34]. In a recent user study, in which thirty four M. Sc. students in management studies were questioned, we found that more than 80% of the students prefer a system that provides document diversity. They indicate that they want document diversity in order to get a more complete map of the information.

Document diversity has many applications. It is considered to be one solution to query term ambiguity. Indeed, queries as expressed by users are not enough to disambiguate terms. To answer such queries, IR can provide the user with a range of documents that corresponds to the various term senses [10]. In that case, redundancy can be penalized by lowering the rank of a document that is too similar to a document already ranked.

Diversity can also be useful in the context of social network analysis [36] and recommender systems [21].

Recommender Systems (RS) aim at providing the user with items related to the current browsed item. Relationships between items can correspond to large range of users' interests and should be linked to their interest in document diversity. Indeed, relevance can be evaluated regarding various criteria. Mothe and Sahut [27] consider the following criteria for information relevance:

- Related to the topic;
- Novelty;
- Understandability;
- Type of media and length;
- Completeness and scope.

In an RS, document similarities are used to associate a score with documents to be recommended. Similarity measures are either based on document content or structure, or on document usage considering popularity or collaborative search.

Similarities from the literature include:

- Similarities based on document content: to be similar two documents should share indexing terms. Example of such measures are the cosine measure [29], or semantic measures [11][32];
- Similarities based on document popularity such as the BlogRank [20];
- Collaborative similarity measures: The document score depends on the scores that previous users assigned to it [1];
- Browsing and classification similarities: document similarity is based either on browsing path [12] or considering the categories users made [5];
- Social similarities based on relationships between contents and users [3][26].

The hypothesis of our work is that diversity in similarity measures is a mean to obtain document diversity for RS. Indeed, not only facets of a topic are diverse but also users are, as well as users' expectations. Even if a unique recommendation method is efficient in the majority of the cases, it is useful to consider the other users' points of views. In that case, diversity on content but also other sorts of diversity should be considered in recommendations. This paper aims at showing the importance of diversity and to present a method and results we obtained when combining various methods in an RS in the context of blogs.

The paper is organized as follows: Section 2 presents the related works. In Section 3, we analyze the overlapping rate of the results retrieved by the best systems that participate to various IR tasks. Section 4 presents a user study on a blog

 ISBN: 978-1-61208-167-0

platform composed of more than 20 million articles. Section 5 concludes this paper.

II. RELATED WORKS

Users are different and therefore their interests are different. To deal with this variety of interests, IR systems try to diversify the retrieved documents. Doing this, they attempt to maximize the chances of retrieving at least one relevant document for the user [30]. However, it is difficult to define what diversity is: several terms are used in the literature to describe this concept.

The literature distinguishes topicality and topical diversity. Topicality refers to which extent the document may be related to a particular topic [33]. Topical diversity groups the extrinsic diversity and the intrinsic diversity. The former helps to dispel the uncertainty resulting from the ambiguity of the user's need or from the lack of knowledge about her needs [28]. The intrinsic diversity, or novelty, intends to avoid redundancy in the retrieved documents [10]. It allows to present to the user various points of view, an overview of the topic that can only be achieved by considering simultaneously several documents, or even to check the information reliability [28].

To introduce topical diversity, two strategies are generally considered either as a clustering problem, or as a selection method close to the Maximal Marginal Relevance (MMR) proposed in [6].

Among the clustering approaches, He *et al.* [15] use Single Pass Clustering (SPC). Better results are obtained in [4] with the k-means algorithm [23]. Assignment to different clusters is generally done by using Euclidean distance and Cosine measure, eventually weighted by the terms frequency. Meij *et al.* [25] apply a hierarchical clustering algorithm on the top fifty documents retrieved by a language modeling approach. The selection of the documents used to build the result list is based on metrics of quality and stability of clusters. The best result from each cluster is selected.

The clustering step usually takes place after a set of documents has been retrieved to reorder them according to the sub-topics identified by the clusters.

Another way to topically diversify the results is to select documents considering those that already occur in the result. To reduce redundancy in the retrieved document list, MMR [6] or sliding window approaches [19] aim at selecting the documents maximizing the similarity with the query and, at the same time, minimizing the similarity with all the documents already selected. The similarity between the document and the previous selection can differ from the similarity with the query [6].

Several approaches select the documents using indicators or filters to increase the diversity in the results. Kaptein *et al.* [19] employ two types of document filters: a filter, which considers the number of new terms brought by the document to the current results and a link filter, which uses the value-added of new input or output links to select new documents. Furthermore, Ziegler *et al.* [37] propose an intra-list similarity metric to estimate the diversity of the recommended list. This metric uses a taxonomy-based classification.

Finally, some user's needs cannot be simply satisfied by topic-related documents. Serendipity, which aims at bringing to the user attractive and surprising documents she might not have otherwise discovered [16], is an alternative to topical diversity. Alternatively, Lathia *et al.* [21] investigate the case of temporal diversity and Cabanac *et al.* [5] consider organizational similarity.

To be able to evaluate and compare topical diversity oriented approaches, TREC Web 2009 campaign defines a dedicated diversity task. This task is based on the ClueWeb09 dataset, which consists in about 25TB of documents in multiple languages. The set B of the corpus, which we use for our experiments, only focuses on the English-language documents, roughly 50 million documents. The diversity task uses the same 50 queries than the *adhoc* tasks (http://trec.nist.gov).

Clarke *et al.* [9] present the panel of metrics used to estimate and compare the performances of the topical diversity approaches. In our experiments, we only consider the Normalized Discounted Cumulative Gain (α-nDCG) [18] and the Intent Aware Average Precision (MAP-IA) [2].

All these approaches and the available evaluation framework are mainly focused on content and on topical diversity. It seems difficult to develop an offline experimental framework, such the TREC Web diversity task, to evaluate other types of diversity, like serendipity. This task is even more difficult in the case of RS. In this context, a users study becomes necessary [14].

III. EXPERIMENTS

We believe that there is no one single approach that would satisfy all users' needs, but a set of complementary approaches. Considering this, we hypothesize that it is interesting to combine them. To demonstrate this, a first step is to verify that two distinct approaches retrieve different documents in various IR contexts (*adhoc*, *diversity*), even if they aim at the same goal.

For this, we consider several systems, which have been evaluated within the same framework to ensure they are comparable. Another selection criterion is the availability of the evaluation runs. Therefore, we focus on the *adhoc* and *diversity* tasks of the TREC Web 2009 campaign (considering only the set B of the corpus). Moreover, we choose the best systems for each task rather than taking into account all the submitted ones.

The runs comparison is done by computing, for each pair of runs, the overlap, that is to say the number of common documents between the two compared runs. The overlap is computed for the N first documents.

In this section, we first compare the four systems, which have the higher MAP at the *adhoc* task and then the four systems with the higher α-nDCG@10 at the diversity task.

A. Adhoc task experiment

1) Adhoc Task and compared runs

"The goal of the task is to return a ranking of the documents in the collection in order of decreasing probability of relevance. The probability of relevance of a document is considered independently of other documents that appear before it in the result list." [8].

The performances of the different systems proposed are compared using the Mean Average Precision (MAP) metric. For this experiment, we consider the best run submitted by the four best competitors at the TREC Web 2009 *adhoc* task. Their scores are presented in Table I.

TABLE I. TREC WEB 2009 ADHOC TASK RESULTS

Group Id	Run Id	MAP
UDel	udelIndDRSP	0,2202
UMD	UMHOOsd	0,2142
uogTr	uogTrdphCEwP	0,2072
EceUdel	UDWAxBL	0,1999

The approaches evaluated (*Run Id*) for these runs are the following ones:

- *udelIndDRSP*: this run combines the query-likelihood language model with the MRF model of term dependencies and the pseudo relevance feedback with relevance models. It also used a document prior called Trust Domain [7];
- *UDWaxQEWeb*: incorporates the semantic term matching method in the axiomatic retrieval framework [35];
- *UMHOOsd*: uses a model based on the Markov Random Fields (MRF) in a distributed retrieval system [22];
- *uogTrdphCEwP*: uses the DPH weighting model derived from the Divergence From Randomness (DFR) model [24].

2) Results

Figure 1 presents the average overlap and precision for the runs selected in the experiments, considering the fifty queries of the task. We note that when considering only first retrieved documents the overlap is low, in spite of the fact that they are most relevant. For example, if we consider the ten first documents for which the precision reaches its highest value (0.386), the overlap is only 22.4%.

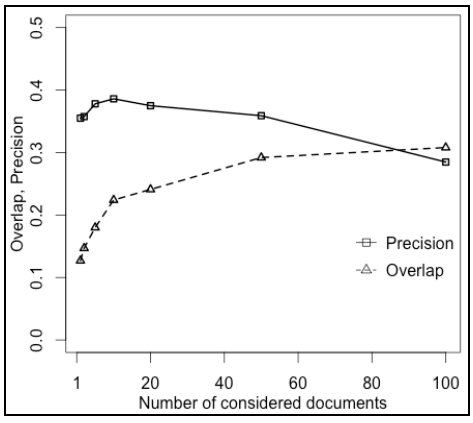


Figure 1. Average overlap and precision for TREC Web 2009 *adhoc* task

These results demonstrate that for a given query, two distinct systems are unlikely to return the same documents. In the next experiment, we check if we get similar observations when we focus on approaches designed to diversify the results.

B. Diversity task experiment

1) Diversity Task and compared runs

In this experiment, similarly to the previous experiment, we center on several systems submitted at the TREC Web Diversity Task. All these systems aim at providing users with diversified result lists.

"The goal of the diversity task is to return a ranked list of pages that together provide complete coverage for a query, while avoiding excessive redundancy in the result list. For this task, the probability of relevance of a document is conditioned on the documents that appear before it in the result list" [8]. The queries are similar for the *adhoc* and the *diversity* tasks. Table II presents the scores obtained by the different systems at their best run.

TABLE II. TREC WEB 2009 DIVERSITY TASK RESULTS

Group Id	Run Id	α-nDCG@10	MAP-IA@10
Waterloo	Uwgym	0.369	0.144
uogTr	uogTrSYCcsB	0.282	0.132
ICTNET	ICTNETDivR3	0.272	0.095
Amsterdam	UamsDancTFb1	0.257	0.082

For the *diversity* task, we retained the following runs:

- *uwgym*: this run acts as a baseline run for the track. It was generated by submitting the queries to one of the major commercial search engines. The results were filtered to keep only the documents included in the set B of the ClueWeb Collection [8];
- *uogTrSyCcsB*: relies upon the DPH DFR model and expands the queries using the Wikipedia documents retrieved [24];
- *ICTNETDivR3*: applies the k-means algorithm to a set of documents using Euclidean distance or Cosine measure to define the nearest cluster [4];
- *UamsDancTFb1*: uses a sliding window approach, which intends to maximize the similarity with the query and, at the same time, to minimize the similarity with the previous document selection. The selection process is completed by a link filter and a term filter [19].

2) Results

As shown in Figure 2, we get similar behavior than in the previous experiment: the overlap is also very low when we consider the first documents. These observations confirm our hypothesis that distinct approaches produce distinct results, even if they attempt to reach the same goal.

C. Conclusion

Whatever the purpose of the different approaches, whether they intend to introduce diversity in the result set, or they are designed to exactly match the needs expressed, the overlap in returned documents is low. Few documents are retrieved in multiple lists. We note that these observations

 ISBN: 978-1-61208-167-0

are especially true when we consider only the first documents, which should theoretically be the most relevant.

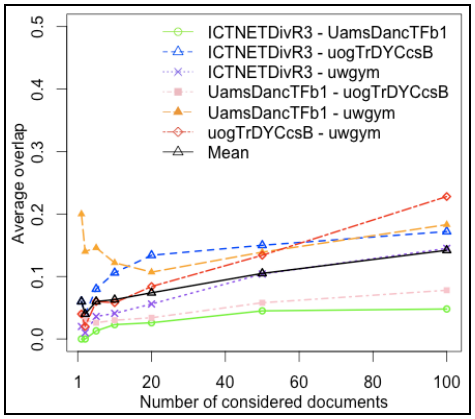


Figure 2. Average overlap for TREC Web 2009 *diversity* task

This leads us to believe that there is no better approach than others. Moreover, it is an indicator that they are complementary. Therefore the choice of the approaches is crucial. Several questions arise:

- How to choose these approaches?
- And how to eventually combine them to maximize the chances of satisfying the user?

IV. USERS STUDY: THE CASE OF OVERBLOG

A. *Diversify the recommendations*

We conducted an experiment with real users to check some hypotheses about the interest of producing diversified recommendations to users in RS. Among the hypotheses are:

- most of the time, users in an IR process search for focus information (topicality);
- sometimes, users want to enlarge the subject they are interested in (topical diversity);
- some users are in a process of discovery and search for new information (serendipity);
- the interesting links between documents do not only concern their content similarity;
- integrating diversity in an RS process is valuable because it allows to answer additional users' needs.

To check these hypotheses, we recruited thirty four students and asked them to test and compare various RS. The users were first asked to type a query on our search engine (first time imposed, to ensure overlap about the documents they all considered, and then a query of their choice). They had then to choose one document and were then shown two lists of recommended documents:

- one list was based on one of the five systems we designed: *mlt* and *searchsim* use topicality, *kmeans* uses topical diversity, and *topcateg* or *blogart* use serendipity (see system description Section B);
- the other list was our RS, designed by merging the results of those five previous systems (choosing the first document in the result list of each system).

Each list contains five documents, and the users do not know which system it corresponds to. They are then asked to choose which list they find the most relevant, and which one they find the most diversified.

Finally, the two lists were mixed into one, and the users has to assess which documents were relevant according to them.

B. *Data and systems*

We focused on the French documents of the *OverBlog* platform. The data used represent more than 20 million articles distributed on 1.2 million blogs.

The five systems used to get recommendation lists are:

- *blogart*: returns documents randomly selected in the same blog of the visited document (serendipity);
- *kmeans*: classifies the retrieved documents with the k-means algorithm [23] and build the final result list selecting one document per cluster (topical diversity);
- *mlt*: uses *Apache Solr's MoreLikeThis* module (http://lucene.apache.org/solr) to retrieve similar documents considering all the content of the visited document (topicality);
- *searchsim*: performs a search using a vector-space search engine. The query used is the title of the visited document (topicality);
- *topcateg*: retrieves the most popular documents randomly selected in the same category (from the OverBlog's hierarchy) of the visited document (serendipity).

The use of these systems aims at simulating the various types of diversity (topicality, topical diversity, serendipity) and intents to limit the overlap between the documents they retrieve. To ensure that the systems used in the user study retrieve distinct results, we compute the overlap between each pair, similarly to the previous experiments. We observe the same trends as in the experiments led on the *adhoc* and *diversity* tasks: the overlap is low between the approaches based on content similarities (*mlt*, *searchsim* and *kmeans*) and is null in the case of serendipity (*blogart*, *topcateg*).

C. *Results*

Table III shows the feedback returned by the user panel concerning the interest of the proposed lists, and their impression on their diversity. For example (4^{th} row), 76.5% of the lists provided by *mlt* have been considered as more relevant than those of *fused*. We can see that the systems perceived as the most relevant are those that focus on topicality. Then the *fused* system is seen as more relevant than its opponent roughly once upon two times on average. We get the same result for *blogart*. That is more surprising, but confirms that users are sometimes interested in links that does not only concern the content of documents. The results obtained for the question "Which one of the following result lists seems the most diversified to you?" is even more surprising, since there are not high differences between the systems. We think this can be explained by the fact that users have difficulties in defining the notion of diversity. We

 ISBN: 978-1-61208-167-0

should have probably helped them by clarifying our question.

TABLE III. PERCENTAGE OF USERS WHO CONSIDER THE *SYSTEM* TO BE MORE RELEVANT/DIVERSIFIED THAN THE *FUSED* SYSTEM

System	Relevance	Diversity
blogart	0.447	0.553
kmeans	0.708	0.333
mlt	0.765	0.500
searchsim	0.643	0.429
topcateg	0.154	0.654

Table IV describes for each RS what is the system precision, that is to say the number of documents that have been considered as relevant among the documents retrieved. We check again that the approaches that use content similarities are seen as more relevant. *kmeans*, that proposes topical diversity, has the best results. On the contrary, *topcateg* and *blogart* that search for serendipity, have lower results. As expected, the *fused* system offers a compromise between these different systems.

TABLE IV. PRECISION PER SYSTEM

System	blogart	kmeans	mlt	searchsim	topcateg	fused
Precision	0.147	0.385	0.265	0.307	0.038	0.267

Finally, Table V compares the fused system with the others. It gives the proportion of relevant documents that have been retrieved by each system. For example, when comparing *mlt* to *fused* (4th column), 54.69% of the relevant documents have been retrieved by *mlt* only, 32.81% by fused only and 12.50% only by both. We can thus observe that, even if more relevant documents come from the systems searching for topicality, a significant part of them comes from the fused system. We think that justifies our approach, because more than 20% of relevant documents coming from our system only means that one document among the five proposed is considered as relevant and would not have been returned if using any system alone.

TABLE V. DISTRIBUTION OF THE RELEVANT DOCUMENTS

System: *fused* against	blogart	kmeans	mlt	searchsim	topcateg
Only retrieved by the system	35.00%	52.46%	54.69%	52.43%	8.77%
Only retrieved by *fused* system	65.00%	21.31%	32.81%	38.83%	91.23%
Commons	0.00%	26.23%	12.50%	8.74%	0.00%

D. Conclusion

The *fused* system we propose offers a new framework to combine various RS. The one implemented and tested here does not outperform the others, but that was not our goal. Rather, our idea is to promote diversity, and we have seen with the user experiments that this is a relevant track. Indeed, by diversifying our recommendations, we are able to answer different and additional users' needs, when the other systems focus on the majority needs: most often the content similarity. The systems we tested here for serendipity were quite simple. Nevertheless, the results they returned were considered as relevant by some users, and we think this is an encouraging sign for developing RS since users are interested in various forms of diversity in result lists.

V. CONCLUSIONS

We are all different, and any system that aims at providing tailored results to their users must take this fact into account. Information retrieval and recommender systems have this ambition, especially since users become used to be given personalized tools, and since they need such systems to overcome the huge amount of data they can access. We have also different expectations depending on the context. Indeed, we may search once for focus information, and the next time search for novelty. This behavior has been confirmed by our user study: the majority of the users' interests has been for topicality (*kmeans, mlt*, *searchsim*), but sometimes, interests for originality (*blogart, topcateg*) emerged. This motivates the design of a system able to handle such heterogeneity of the users' needs.

Our first contribution in this area has been to study the overlap between the documents retrieved by several IR systems of the literature using state-of-the-art datasets. Although those systems are all content similarity-based, we have noted that they are based on different underlying assumptions, and that their overlap was low. This low overlap between the documents retrieved by these systems indicates that there is not a perfect system able to satisfy the diversity of the users' needs, but a set of complementary systems.

The strength of our approach is that it is designed to combine different RS. We could easily add to the framework any other RS that would cover new interests. It seems that some work could especially be done on the approaches that offer serendipity: *blogart* seems to be an interesting one; *topcateg* is to be improved. Other RS fusion approaches have been proposed in the literature. In particular, Shafer *et al.* [31] and Jahrer *et al.* [17] present a "meta RS". However, when they choose to focus on results shared by the different RS used by the meta system, we instead propose to select the best recommendations of each system to ensure diversity. We assume that it is important to give a chance to all possible "points of view" proposed by every retrieved document.

We will direct our future work towards designing an RS architecture promoting the diversity of recommendations. When existing approaches focus on designing methods to force diversity in their results (using clustering or MMR), we choose to consider multiple systems to build the recommendation list and ensure diversity. Moreover, it is important to see that every retrieved document may give rise to a wide range of interests for readers. So, next step for this work is to study the learning mechanism to find the proportion of documents coming from each RS to be fused. That is to say we would learn the main interests that are important for end-users (readers). To do this, our idea is to use an automatic learning process based on the users' feedbacks. We could for example simply initialize the system with equal distribution for each RS, and then increase the proportion of recommendations for the ones that are more often clicked by the users, and decrease the proportion

 ISBN: 978-1-61208-167-0

for RS less often considered. Our meta RS would thus be well focused on the users' needs. Considering the results of the experiments presented in this paper, we could expect a 80% proportion for topicality systems, and 20% for more original systems.

Finally, we will see if these assumptions stand on a real scale experiment using the online blog platform *OverBlog*. Then, we will conduct statistical analysis to study what could be the influence of the document type on those proportions.

REFERENCES

[1] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions", IEEE Transactions on Knowledge and Data Engineering, 17(6), 2005, pp. 734-749

[2] R. Agrawal, S. Gollapudi, A. Halverson, and S. Ieong, "Diversifying search results", Inter. Conf. on Web Search and Data Mining, 2009, pp. 5-14

[3] L. Ben Jabeur, L. Tamine, and M. Boughanem, "A social model for Literature Access: Towards a weighted social network of authors", Inter. Conf. on Adaptivity, Personalization and Fusion of Heterogeneous Information RIAO, 2010, pp. 32-39

[4] W. Bi, X. Yu, Y. Liu, F. Guan, Z. Peng, H. Xu, and X. Cheng, "ICTNET at Web Track 2009 diversity task", Text REtrieval Conf., 2009

[5] G. Cabanac, M. Chevalier, C. Chrisment, and C. Julien, "An Original Usage-based Metrics for Building a Unified View of Corporate Documents", Inter. Conf. on Database and Expert Systems Applications, 2007, LNCS V. 4653, 2007, pp. 202–212

[6] J. Carbonell and J. Goldstein, "The use of MMR, diversity-based reranking for reordering documents and producing summaries", ACM Conf. on Research and Development in Information Retrieval, 1998, pp. 335-336

[7] P. Chandar, A. Kailasam, D. Muppaneni, L. Thota, and B. Carterette, "Ad Hoc and Diversity Retrieval at the University of Delaware", Text REtrieval Conf., 2009

[8] C. L. A. Clarke, N. Craswell, and I. Soboroff, "Overview of the TREC 2009 Web Track", Text REtrieval Conf., 2009

[9] C. L. A. Clarke, N. Craswell, I. Soboroff, and A. Ashkan, "A comparative analysis of cascade measures for novelty and diversity", ACM Inter. Conf. on Web Search and Data Mining, 2011

[10] C. L. A. Clarke, M. Kolla, G. V. Cormack, O. Vechtomova, A. Ashkan, S. Büttcher, and I.n MacKinnon, "Novelty and Diversity in Information Retrieval Evaluation", ACM Conf. on Research and Development in Information Retrieval, 2008, pp. 659-666

[11] D. Dudognon, G. Hubert, J. Marco, J. Mothe, B. Ralalason, J. Thomas, A. Reymonet, H. Maurel, M. Mbarki, P. Laublet, and V. Roux, "DYNAMic ontology for information retrieval", Inter. Conf. on Adaptivity, Personalization and Fusion of Heterogeneous Information RIAO, 2010, pp. 213-215

[12] I. Esslimani, A. Brun, and A. Boyer, "A Collaborative Filtering approach combining Clustering and Navigational based correlations", Web Information Systems and Technologies, 2009, pp. 364-369

[13] E. A Fox and J. A. Shaw, "Combination of multiple searches", Text REtrieval Conf., NIST special publication, 1994, pp. 243–252

[14] C. Hayes, P. Massa, P. Avesani, and P. Cunningham, « An online evaluation framework for recommender systems », Workshop on Personalization and Recommendation in E-Commerce, 2002

[15] J. He, K. Balog, K. Hofmann, E. Meij, M. de Rijke, M. Tsagkias, and W. Weerkamp, "Heuristic Ranking and Diversification of Web Documents", Text REtrieval Conf., 2010

[16] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating Collaborative Filtering Recommender Systems", ACM Trans. Information Systems, 22(1), 2004, pp. 5-53

[17] M. Jahrer, A. Töscher, and R. Legenstein, "Combining predictions for accurate recommender systems", ACM SIGKDD Inter. Conf. on Knowledge Discovery and Data mining, 2010

[18] K. Järvelin, and J. Kekäläinen, "Cumulated gain-based evaluation of IR techniques", ACM Transactions on Information Systems, 20(4), 2002, pp. 422–446

[19] R. Kaptein, M. Koolen, and J. Kamps, "Result diversity and entity ranking experiments: Anchors, links, text and Wikipedia", Text REtrieval Conf., 2009

[20] A. Kritikopoulos, M. Sideri, and I. Varlamis, "Blogrank: ranking weblogs based on connectivity and similarity features", Inter. Workshop on Advanced Architectures and Algorithms for Internet Delivery and Applications, 2006, art. 8

[21] N. Lathia, S. Hailes, L. Capra, and X. Amatriain, "Temporal diversity in recommender systems", ACM Conf. on Research and Development in Information Retrieval, 2010, pp. 210-217

[22] J. Lin, D. Metzler, T. Elsayed, and L. Wang, "Of Ivoryand Smurfs: Loxodontan MapReduce Experiments for Web Search", Text REtrieval Conf., 2009

[23] J. MacQueen, "Some methods for classification and analysis of multivariate observations", Berkeley Symposium on Mathematical Statistics and Probability, 1967, pp. 281–297

[24] R. McCreadie, C. Macdonald, I. Ounis, J. Peng, and R. L. T. Santos, "University of Glasgow at TREC 2009: Experiments with Terrier", Text REtrieval Conf., 2009

[25] E. Meij, J. He, W. Weerkamp, and M. de Rijke, "Topical Diversity and Relevance Feedback", Text REtrieval Conf., 2010

[26] J. Mothe, C. Chrisment, T. Dkaki, B. Dousset, and S. Karouach, "Combining mining and visualization tools to discover the geographic structure of a domain", Computers, Environment and Urban Systems, 30(4), 2006, pp. 460-484

[27] J. Mothe and G. Sahut, "Is a relevant piece of information a valid one? Teaching critical evaluation of online information". Approaches to Teaching and Learning Information Retrieval, N. Efthimiadis, Manuel Fernández Luna, Juan Huete, Andrew MacFarlane (Eds.), 2011 (to appear)

[28] F. Radlinski, P. N. Bennett, B. Carterette, and T. Joachims. "Redundancy, diversity and interdependent document relevance", SIGIR Forum, 43(2), 2009, pp. 46–52

[29] G. Salton and M. J. McGill, "Introduction to Modern Information Retrieval", McGraw-Hill, 1983

[30] R. L. T. Santos, C. Macdonald, and I. Ounis, "Selectively Diversifying Web Search Results", ACM Inter. Conf. on Information and Knowledge Management, 2010

[31] J. B. Schafer, J. A. Konstan, and J. Riedl, "Meta-recommendation systems: user-controlled integration of diverse recommendations", Inter. Conf. on Information And Knowledge Management, 2002, pp. 43-51

[32] Z. Wu, M. Palmer, "Verb semantics and lexical selection", Annual Meeting of the Associations for Computational Linguistics, 1994, pp. 133-138

[33] Y. C. Xu and Z. Chen, "Relevance judgment: What do information users consider beyond topicality", Journal of the American Society for Information Science and Technology, 57(7), 2006, pp. 961–973

[34] X. Yin, J. X. Huang, and Z. Li, "Promoting Ranking Diversity for Biomedical Information Retrieval Using Wikipedia", European Conf. on Information Retrieval, 2010, pp. 495-507

[35] W. Zheng and H. Fang, "Axiomatic Approaches to Information Retrieval - University of Delaware at TREC 2009 Million Query and Web Tracks", Text REtrieval Conf., 2009

[36] X. Zhu, A. B. Goldberg, J. Van Gael, and D. Andrzejewski, "Improving Diversity in Ranking using Absorbing Random Walks", Human Language Technologies - Conf. North American Chapter Assoc. Comp. Linguistics, 2007, pp. 97-104

[37] C. Ziegler, S. McNee, J. A. Konstan, and G. Lausen, "Improving recommendation lists through topic diversification", Inter. Conf. on World Wide Web, 2005, pp. 22–32

 ISBN: 978-1-61208-167-0

A Method to Synthesize 3-Dimensional Face Corresponding to Diverse Words Expressing Facial Features through Mapping and Inspecting Validity of Mapping Function

Futoshi Sugimoto
dept. of Information Sciences and Arts
Yoyo University
Kawagoe, Japan
e-mail: f_sugi@toyo.jp

Abstract— **In this study, the process to synthesize a human face based on the information of words is defined as a mapping from a word space, which is composed of the words expressing dimensions and shape of facial elements, into a physical model space where physical shape of the facial elements are formed. By introducing a concept of mapping, the use of whole words existing in the word space makes it possible to synthesize a human face based on free and uninhibited description. Furthermore, we have only to make 3-dimensinal physical models corresponding to the words that are selected as training data to identify a mapping function. The others are made through the mapping. Finally, we inspect the validity of the mapping function that is obtained in this study.**

Keywords- three dimensional facial synthesis; information of words; mapping function; computer graphics; GMDH.

I. INTRODUCTION

In recent years, since the importance of the facial information has been recognized, various studies and the practical use to computerize the information have been performed, and many papers and articles on them were published [1][2][3]. These researches can be divided into recognition and synthesis from the viewpoint of information processing. Our study is set in the area of the face synthesis among these researches. We aim at constructing a system to synthesize a 3-dimensional face by using computer graphics based on the information of words, which can describe facial features very freely (we call them "feature words" in this study). In this paper, we are to propose a method to form physical models of facial elements corresponding to the feature words using a mapping function, which plays a main role in our face synthesis system, and then we are also to inspect the effectiveness of the mapping function.

When we try to describe facial features of a person whom we picture to ourselves, the description is by using several distinct levels of words. Some words may describe directly and concretely the physical dimension and the shape of facial elements; while others may do so abstractly or metaphorically. In former studies on synthesis of a human face by utilizing the information of words, only few words describing directly the physical features were used, and the words explaining some degree of physical dimension, i.e., "slight," "a little," "very," and so on, were simply added to [4].

In this study, our main aim is to synthesize a face based on free and uninhibited description. Which means we can use abstract and metaphorical words as well as concrete and physical ones as the feature words. In order to realize it, we define the process to synthesize a human face based on the information of words as a mapping from a word space to a physical model space. The word space and the physical model space are to be explained in another section in detail.

Introducing the concept of mapping enables us to synthesize the 3-dimensional physical models corresponding to diverse words [5][6][7][8]. We adopted GMDH (Group Method of Data Handling) [9] to identify a mapping function in this method. It is because the effectiveness of GMDH to identify a mapping function under the conditions whose relations are complicated and non-linear, and there are little training data, is already verified [10][11]. In this paper, we focus on inspecting the usefulness of the mapping and the validity of the mapping function that is obtained by GMDH.

The contents of the paper is as follows; In Section 2, the system we have been developing is outlined. In Section 3, the process to construct the word space and its characteristics are described. In Section 4, the physical model space is described. In Section 5, the process to make the training data and to identify the mapping function is described. In Section 6, the validity of the mapping function that is obtained in this study is inspected. Finally, in Section 7, the conclusion and the future works are presented.

II. OUTLINE OF THE SYSTEM

The outline of the facial synthesis system that we have been developing is shown in Figure 1. This system has a word space and a physical model space, which are to be explained in another sections in detail, and the process to synthesize a physical model of a human face is defined as a mapping from the word space to the physical model space. The facial elements in this research are nose, eyes, mouth, eyebrows, cheeks, jaw, and profile. The word space and the physical model space are made for each facial element, and the mapping is executed for each facial element respectively.

 ISBN: 978-1-61208-167-0

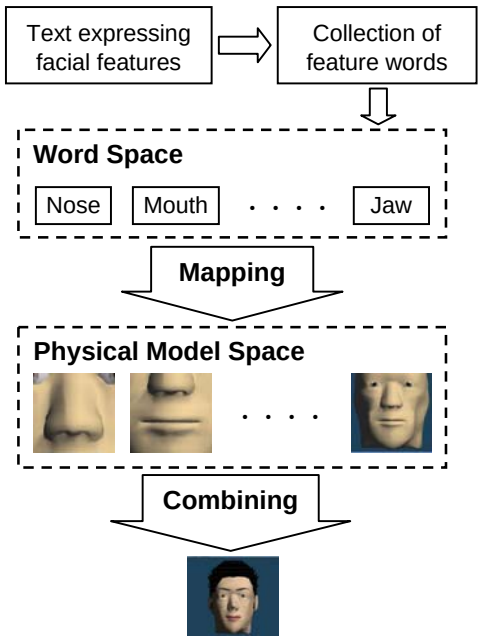


Figure 1. Outline of the system to synthesize 3-dimensional face

Before synthesizing facial elements, the feature words are collected from a sentence describing facial features or testimony of a witness, which is not, however, included in our current research. A physical model corresponding to an extracted feature word is made through mapping every individual facial element, and then a human face is synthesized through combining all physical models of facial elements together. This paper focuses on the part of making the physical models of facial elements by mapping.

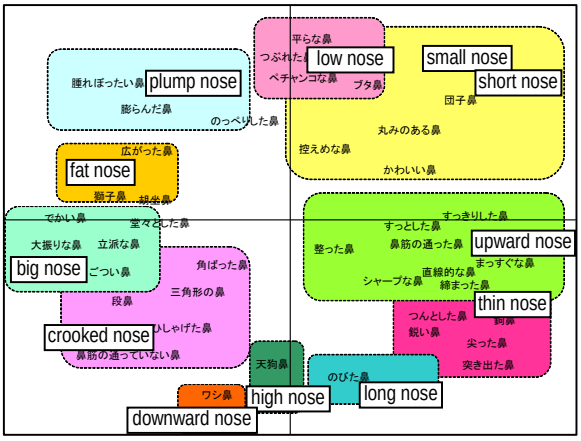


Figure 2. Word space, clustering result and training data of feature words in case of nose

III. WORD SPACE

The word space in this study is composed of the feature words, which express dimensions and shape of facial elements, and it is constructed for each facial element.

A. Construction of Word Space

In order to construct the word space, firstly many feature words were collected for each individual facial element, and secondly those words were located in a space by Multi-Dimensional Scaling method (MDS) [14] based on the similarity of those feature words. We call this space the word space. Those feature words were from a Japanese dictionary [12] and [13]. We provided four following criteria on collecting the feature words.

(1) The nouns that express figuratively the feature of the facial elements.
(2) The adjectives that express the feature of the facial elements when they are put before a facial element.
(3) The mimetic words that express the feature of the facial elements.
(4) The words that are used for expressing the feature of the facial elements in our daily conversation.

A similarity matrix among the feature words to be the input data of MDS was obtained using a method [15] of applying information theory in this study. In this method, the subjects (50 male students in our department, around 22 years old) classified the feature words based on the similarity of the impression with which the feature words are associated, and probability that the feature words are classified in the same group was found, and finally similarity among the feature words was calculated based on the probability. In the word space obtained from MDS, the more similar the feature words are, the closer they are located, while the farther, the less similar.

Every word space has six dimensions in this study. This is determined based on an indicator called “stress,” which shows how the distance relationship in the word space satisfies the similarity relationship among the feature words. Since a feature word is a point in the 6-dimensional word space, a feature word in the word space of a facial element is defined as $\mathbf{W}_i$, and it is described as follows;

$$\mathbf{W}_i = (w_1, w_2, ..., w_6), \quad i = 1, ..., m \qquad (1)$$

Here, m is the number of the feature words in a word space, w_j expresses coordinate value of jth axis in the 6-dimensional space.

B. Characteristics of Word Space

The word space of nose is shown in Figure 2 as an example. Since we cannot find any proper English words corresponding to all Japanese feature words, we display the word space in Japanese except the training data. Although the word space is 6-dimensional in reality, it is projected on a two dimensional plane for visually understanding. The

 ISBN: 978-1-61208-167-0

characteristics of the word space summarized to be seen in Figure 2 are as follows;

(1) The feature words that have completely opposite meanings stand face to face each other across the origin of coordinates.
(2) Almost the feature words tend to be located at the edge of the word space.
(3) The feature words that have the meaning of almost a standard feature are located near the center.

By the analogy from the characteristics mentioned above, it is appropriate to think that the "standard feature" is located at the origin of the word space, and the feature words that have an adjective that expresses the degree of feature such as "slight" and "very" are located on the straight line connecting the origin and a certain feature word.

IV. PHYSICAL MODEL SPACE

The physical model space in this study also is constructed for each facial element. It is composed of the physical shape of the facial element corresponding to each feature word.

A. Construction of Physical Model Space

The 3-dimensional geometric model of facial element corresponding to each feature word is made as a wire frame model by computer graphics (CG). In this study, the wire frame model is called the physical model of the feature words, and the space composed of the physical models is defined as physical model space. A physical model $\mathbf{M}_i$ corresponding to a feature word $\mathbf{W}_i$ of a facial element is a set of apexes of the wire frame model, which is described as follows;

$$\mathbf{M}_i = (\mathbf{P}_{i1}, \mathbf{P}_{i2}, ..., \mathbf{P}_{in}) \quad (2)$$

Here, n is the number of apexes of the wire frame model for each facial element. $\mathbf{P}_{ij}$ is jth apex of the wire frame model, and it is composed of xyz coordinates as shown in (3).

$$\mathbf{P}_{ij} = (x_{ij}, y_{ij}, z_{ij}), \quad j = 1, ..., n \quad (3)$$

Since the number of apexes is different from each facial element, the physical model space for each facial element has a different dimension (3 x the number of apexes) from each other.

B. Design of Standard Face Model

The standard face model used in this study is a Japanese man who is about 22 years old. The process to make the standard face model is as follows; First of all, the photographs of the face of 40 male university students were taken from the front and the side, and 34 items (some examples shown in Figure 3) were measured, and then the mean values of the items were calculated. Secondly, a wire frame model of the standard face was designed based on the mean values. Finally, the standard face model was completed by mapping the textures of eyes, eyebrows, lips and skin that are shown in Figure 4 on the wire frame model.

The standard face model is divided into 5 facial elements, e.g. nose, eyes, mouth, cheeks, and jaw, as shown in Figure 5, and their shape can be deformed for each element. The

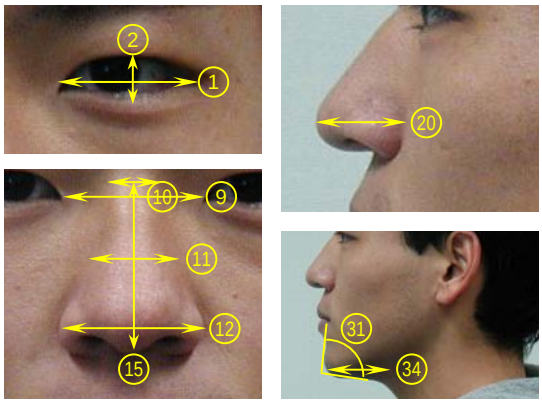


Figure 3. Examples of measurement items

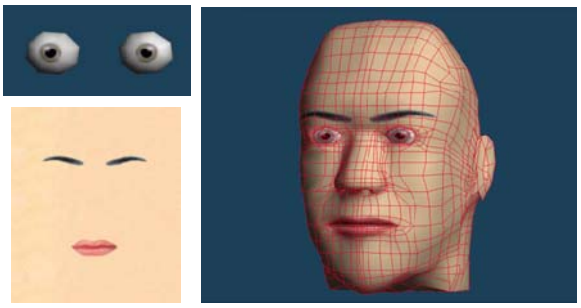


Figure 4. Texture and standard face model

Figure 5. Division into five facial elements

 ISBN: 978-1-61208-167-0

eyebrows are deformed by changing their shape, position and leaning of the texture, not by 3-dimensional model. It is not necessary to make the physical models for all feature words by manual labor for each facial element. Only the physical models for the feature words chosen for training data that is explained in Section 5 are needed to be made. Concerning the model of another feature words except training data, the coordinates of apexes of wire frame model are calculated by a mapping function, and the wire frame model becomes the physical model of other feature words.

V. MAPPING FUNCTION

There are many feature words in the word space of a facial element. In order to identify the mapping function, we need to select several training data from the feature words and to make physical models corresponding to the selected feature words.

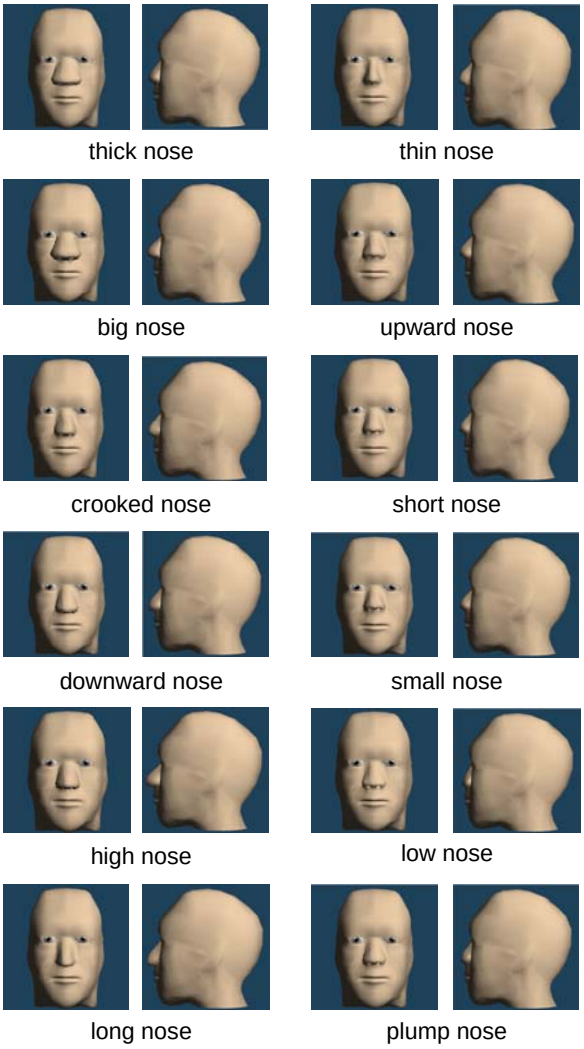


Figure 6. Physical models of nose corresponding to each training data of feature word

A. Training Data

It is necessary that several feature words are extracted for training data from the word space equally in space for each individual facial element respectively. At first, the feature words were classified using cluster analysis based on Euclid distance among the feature words in the word space. Next, the representative was selected from each cluster, and they became the training data. The clusters of the feature words in the case of nose are shown classifying with colored area in Figure 2, and the words that are selected for the training data are enclosed with a square. Only the training data words are displayed in English for the reason mentioned in Section 3.

The training data in the physical model space corresponding to the one in the word space is necessary to be made. Several photographs, which have the facial element having the impression with which the training data is associated were picked up from 40 photographs mentioned in Sections 4. The physical models of the training data were made by manual labor based on the average value of the measured items of the selected photographs. Figure 6 shows the twelve training data words of nose and the physical models corresponding to the words.

B. Identification of Mapping Function

A set of *xyz* coordinates of all the apexes in the wire frame model becomes the parameters of the physical model space. We identify the mapping function from the training data using a statistical method, GMDH. The mapping function can be described as follows;

$$\mathbf{M}_i = \mathbf{f}(\mathbf{W}_i) \tag{4}$$

Since a physical model $\mathbf{M}_i$ is a set of apexes of the wire frame model as shown in equation (2), the mapping function for each apex becomes as follows;

$$\mathbf{P}_{ij} = \mathbf{f}_j(\mathbf{W}_i) \tag{5}$$

Furthermore, since an apex $\mathbf{P}_{ij}$ is composed of *xyz* coordinates as shown in equation (3), the actual mapping function becomes as follows;

$$x_{ij} = f_{ij}(\mathbf{W}_i), y_{ij} = f_{yj}(\mathbf{W}_i), z_{ij} = f_{zj}(\mathbf{W}_i) \tag{6}$$

A set of functions are obtained for each individual facial element respectively. The number of mapping function for each facial element is 3 x the number of apexes of the wire frame model.

VI. INSPECTION OF MAPPING FUNCTION

We inspect the validity of the physical models, which are made by this system using the mapping function in this section. Therefore, the questionnaire including 36 sets such as shown in Figure 7 was presented to 20 subjects, and they

 ISBN: 978-1-61208-167-0

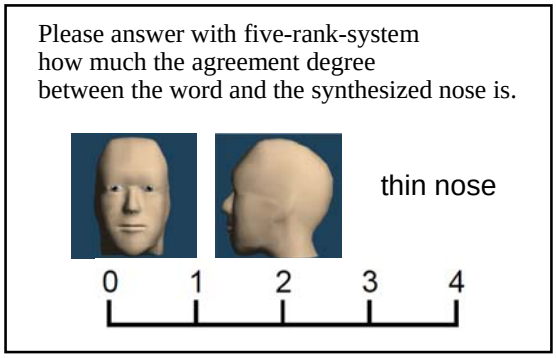


Figure 7. An example of questionnaire used in experiment to inspect validity of physical models

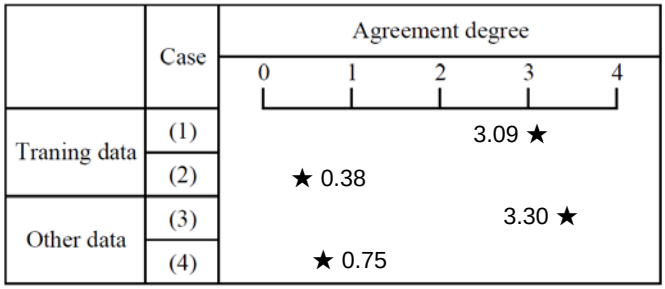


Figure 8. Result of experiment to inspect validity of physical models

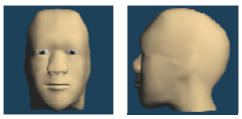
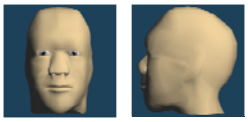
were required to evaluate the agreement degree between the feature word and the physical model with five-rank-system. The subjects are male students in our department, and are around 22 years old. The 36 sets that combine the feature word and the physical model are as follows;

(1) The training data model and the feature word corresponding to it, 12 sets.
(2) The training data model and the feature word having opposite meaning, 6 sets.
(3) The model belonging to the same group of the training data and the feature word corresponding to it, 12sets.
(4) The model belonging to the same group of the training data and the feature word having opposite meaning, 6 sets.

(2) and (4) are inserted into the questionnaire in order to check the reliability of the subjects. The examples that used in each case in the experiment are shown in Table I.

The result of the questionnaire is shown in Figure 8. The evaluation of case (1) is slightly lower than case (3). The reason is that since the words directly express physical shape, they are selected as the training data, so it seems that the subjects may expect a more typical shape. On the other hand,

TABLE I. EXAMPLES USED EACH CASE IN THE EXPERIMENT

Case	Feature word	Physical model
(1)	thin nose	
	thick nose	
(2)	thick nose	
	small nose	
(3)	flattened nose	
	tight nose	
(4)	pointed nose	
	snub nose	

since the other words except the training data are abstract, the subjects tend to easily agree. The portrait resembles the person himself if the characterized part is emphasized more than necessary. In the same way as this effect, the training data model may be the extreme geometric model that the part is emphasized more in our system too. However, it may be said that the purpose to make the physical models corresponding to the feature words except training data using mapping function can be accomplished enough.

VII. CONCLUSION AND FUTURE WORK

We propose a method to synthesize a 3-dimensional face from the information of words in this paper. This method allows a user to use the words abstractly or figuratively expressing the physical shape of facial elements as well as the words directly expressing it. The characteristics of this method is that it defines the process where a human face is synthesized based on the information of words as a mapping

 ISBN: 978-1-61208-167-0

from the word space to the physical model space. Using this method, it becomes possible to synthesize a human face corresponding to all words in the word space.

Finally, we describe the future challenge and prospect. The process to synthesize human 3-dimensional face by combining the physical models of facial elements together is already completed, so we will perform the evaluation immediately in the future and publish it in an article. Although the current standard face model is a Japanese man who is about 22 years old, we will make it for each sex, age and for typical shape of facial profile, and then we will make it possible to synthesize more various face models.

REFERENCES

[1] M. Kaneko Ed., "Now, How Interesting a "Face" Is!; Processing of Facial Images and Its Application," The Journal of the Institute of Image Information and Television Engineers, Vol. 62, No. 12, pp. 13-41, 2008. (in Japanese)

[2] R. Valenti, A. Jaimes and N. Sebe, "Facial Expression Recognition as a Creative Interface," Proceedings of 2008 International Conference on Intelligent User Interfaces, pp. 433-434, 2008.

[3] S. Kumano, K. Otsuka K, J. Yamato, E. Maeda, and Y. Sato, "Pose-Invariant Facial Expression Recognition using Variable-Intensity Templates," International Journal of Computer Vision, Vol. 83, No. 2, pp. 178-194, 2009.

[4] S. Iwashita and T. Onisawa, "Facial Caricature Drawing with Personal Impressions," Proceedings of 5th International Conference on Soft Computing, pp. 209-212, 1998.

[5] Y. Hoshino and F. Sugimoto F, "Forming 3-Dimensional Face Model Based on the Information of Words Expressing Facial Feature," Proceedings of the 68th Annual Convention IPS Japan, Vol. 4, pp. 69-70, 2006. (in Japanese)

[6] K. Mochida and F. Sugimoto, "A System to synthesize 3-Dimensional Face Based on the Flexible Verbal Expression," Proceedings of the 2006 Symposium of Human Interface Society, pp. 1139-1142, 2006. (in Japanese)

[7] F. Sugimoto and M. Yoneyama, "3D Face Synthesis Based on the Information of Words Expressing Facial Features," Proceedings of IEEE Workshop on Computational Intelligence in Virtual Environments, pp. 1-6, 2009.

[8] F. Sugimoto, "A Method to Visualize Information of Words Expressing Facial Features," Proceedings of IEEE Computer Society 2010 Fifth International Multi-conference on Computing in the Global Information Technology, pp. 169-174, 2010.

[9] A. G. Ivakhnenko, "Polynomial theory of complex systems," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-1, No. 4, pp. 364-378, 1971.

[10] N. Honda, F. Sugimoto, and S. Aida, "Analysis of Cartoon Faces and Design of Face Pattern Used in Faces Method," The Bulletin of the University of Electro-Communications, Vol. 31, No. 1, pp. 1-10, 1980. (in Japanese)

[11] A. G. Liao, W. F. Liu, and G. Y. Xiang, "The prediction research on the vehicles for business transport of Guangxi in China with the GMDH algorithm method," Proceedings of 2010 International Conference on Computer Application and System Modeling, Vol. 3, pp. 36-38, 2010.

[12] K. Kindaichi, Japanese Dictionary, the Fifth Edition, Sanseido, 2001. (in Japanese)

[13] R. Konomiya Ed., Dictionary to Explain Body, Shintensya , 2002. (in Japanese)

[14] C. Hayashi A. and Akuto, Multi-Dimensional Scaling Method, pp. 3-160, Science Co., 1976. (in Japanese)

[15] T. Saito, Multi-Dimensional Scaling Method, pp. 197-200, Asakurasyoten, 1980. (in Japanese)

 ISBN: 978-1-61208-167-0

User Profile Matching: A Statistical Approach

Work in Progress Paper

Benjamin Wassermann, Gottfried Zimmermann
Wilhelm Schickard Institute of Computer Science
Eberhard Karls University of Tübingen
Germany
benjamin.wassermann@uni-tuebingen.de
gottfried.zimmermann@uni-tuebingen.de

***Abstract* - Users interact with many applications and user interfaces, optionally involving additional user interface components such as assistive technologies. Adaptations of user interfaces, including the activation of assistive technologies, typically require the user to perform manual adaptations for every new device and application. It would be beneficial for the user, if adaptation data could be automatically collected and stored in a standardized way, to be utilized by any device and application the user interacts with. In this paper we present a new approach for user interface adaptation, based on statistical methods. These methods operate on an extended set of user interaction-related properties, which may be part of user profiles. We plan to implement this approach in a prototype on real user data and evaluate its results against manual adaptations by user interface experts.**

Keywords - Adaptive user interface, user interface matching, user profile, user profile matching, statistical analysis.

I. INTRODUCTION

The Internet is a fast evolving network exposing new technologies in a fast pace. In one of the current trends, user-generated files and data are moved from the local desktop to the cloud in the Web. This means that the user stores their documents, emails, pictures, etc. on Internet-based servers rather than their personal computers. Not only is the data stored on the Internet; the application itself to create and edit the data is provided over the Internet, running directly in the Web browser. This allows for accessing and editing the same data on a broad range of devices (personal computer, mobile phones, etc.) that have access to the Internet all over the world.

Currently, many new and versatile Web applications are emerging in the Internet. The user is confronted with new dynamic user interfaces on various devices that differ from classical Websites and local applications. These new Web applications give rise to new opportunities in user interaction, in that dynamic user interfaces can be adapted not only for the device and environmental contexts, but also for the needs and preferences of any user. Such user interface adaptations have the potential to increase the usability of Web applications in general, and particularly for people with disabilities.

We identify three different approaches of automatic user interface adaptations: The rule based approach uses heuristics to make decisions on adaptations; the statistical approach uses large collections of recorded user data to make decisions based on statistical analysis of the data. A third approach combines both previous approaches as a hybrid.

Statistical approaches for matchmaking are used in different fields of applications. One important application field is image and video processing. Statistical algorithms are used to find patterns and objects, improve quality, predict movement, etc. [1].

On the Web, statistical approaches are currently being used for various purposes:

- Amazon is collecting data from the users while they are using their shopping portal. Amazon uses statistical algorithms on the collected data to find user preferences on products and align this information with similar users to create selective shopping suggestions for a user [2].
- Dating agencies are using statistical methods to find good matches between people. They analyze the user profiles and their collected data about successful meetings to continuously improve their matching [3].

While these applications use statistical approaches for making recommendations to the user, we want to use statistical algorithms for user interface adaptations based on user profiles.

Research on user interface adaptations is being conducted in a number of projects, but none of them uses a statistical matching approach. The European FP7 projects MyUI [4] and GUIDE [5] are examples of such projects.

Before we describe our concept, we will set the scene with a scenario that describes an instance of what the concept of this paper could achieve. After that we will give a brief introduction to the technical backgrounds for user profiles, user profile matching and user interface matching based on user profiles. In section 4 we will discuss the evaluation methods we intend to use to prove the concepts. Finally, we give an outlook for how we plan to implement and validate the concepts of this paper.

 ISBN: 978-1-61208-167-0

II. EXAMPLE SCENARIO

Martha (31 years old) lives alone in her flat in the city center. She has a notebook for surfing in the Web and for writing emails. She also has a smartphone, but uses it only to make calls or to write text messages. At her work, she has a personal computer and writes product reviews.

Martha uses an office suite for writing texts, both at home and at work. She is glad to have managed to use this application, and doesn't know much about other software. To perform the functions in her office suite, she uses the computer mouse in the office and the touchpad on her notebook at home. She doesn't use keyboard shortcuts, except for copy and paste. For other features, such as text formatting, she uses the toolbar of the office application.

At work she got an assignment to write a report, together with her colleagues. They have agreed to use a new Internet-based tool that allows them to work in parallel on the same text. It also promises to adapt automatically to the user's needs and preferences.

Martha goes to the website and registers for the new tool, so that she can use it. Before she can start, she has to fill out a questionnaire, asking a few questions about her preferences for different applications, such as what software she uses, whether she uses shortcuts and, if so, what kind of shortcuts.

In fact, the user interface of the Web application looks very similar to her desktop program. She can start working immediately without learning the new interface. There are a few functions she doesn't know. However, the program informs her about the new functionality and tells her that other users (that have a similar profile to hers) have rated these functions as very practical. Martha reads the descriptions and decides that two of the four functions might be useful for her and disables the other two functions. Internally, the system recognizes the user adaptations and stores them in the user profile, so the functions will no longer appear.

Over the time, Martha starts to use other Web applications and is very pleased how well the user interfaces adapt to her behavior and needs. In particular, changes done in one application are also reflected in the other applications. Meanwhile, she wants this functionality for all of her applications.

III. TECHNICAL BACKGROUND

A. User Profile

The foundation for adapting a user interface to a specific user in general, and in particular for above scenario, is the user profile. It stores various data that can be used to derive an optimal adaptation of the user interface.

Data mining of user profiles is essential if we want to apply statistical methods to user interface adaptation. We need to operate on rich information items on the user, which is more than mere demographic data. Information about the user preferences and patterns of typical user interactions are required to make informed and reasonable decisions. This can be achieved by user modeling techniques [6] [7].

A user profile should therefore include:

- Demographic data;
- Preferences on features, settings, shortcuts, etc.;
- Preferred user interaction strategies (user interaction patterns);
- User interaction enhancements, such as assistive technologies.

Of course, as the profile holds sensitive user data, security and privacy issues about data storing and data access need to be thoroughly considered in an actual implementation. However, in this paper we focus on the concept of user profile matching.

B. User Profile Space

An initial user profile is normally incomplete, because user interaction patterns or user preferences are recorded gradually while the user works with various applications. In general, only demographic data and simple preference data, which is present at the beginning, is collected by an initial questionnaire. This data is not sufficient to produce any reliable results by the means of statistical analysis. One possible solution to this problem is to borrow the missing data from other similar, though more complete, user profiles. For this task a user profile space is needed, that is generated from a high number of user profiles. Such profiles are collected over a long period of time to yield enough interactions and preferences of the users. It is also possible to import user profiles from existing profile databases or to create them from virtual users or personas.

A new user profile is classified by a profile matching component and is afterwards integrated into the user profile space, using the available data stored in the profile (see fig. 1). After the classification, similar user profiles can be used to fill the gaps in the new user profile. It is important to note that at this point, this data is only a hypothesis by the system, and is therefore assigned with appropriate probability levels. However, with every new interaction of the user with any application, this hypothetical information becomes either corrected or enforced, resulting in higher probability scores.

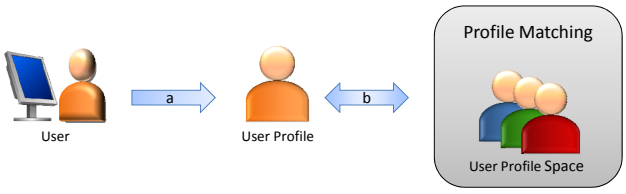


Figure 1. Sequence for (a) creating a new user profile and (b) matching it to the user profile space. After the matching the user profile is supplemented with data from similar profiles.

By following this process, we can constantly extend and improve the user profile space, with new user profiles being integrated and existing profiles validated by incoming user data (see fig. 2).

We will use a statistical approach for matching user profiles into the user profile space. One candidate is the principal component analysis (PCA) [8]. This method transforms a high dimensional feature space, derived from the user profiles, to so-called principal components. Ideally,

 ISBN: 978-1-61208-167-0

no data gets lost with the transformation, and redundant data gets combined to correlated data.

A new user profile can be aligned along the principal components to get an optimal classification in the user profile space. Now, the correlated data can be used for the completion of the missing information of a user profile. Additionally, a plausibility weight in form of a probability is assigned for any derived information item of the user space. This probability is continuously adjusted to user profile space updates, e.g., when new user profiles are added or profiles get updated with new data from a user.

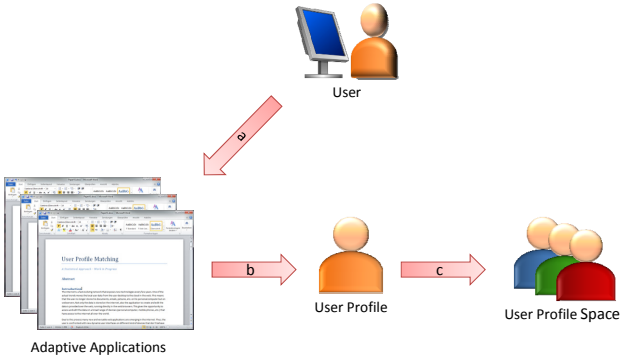


Figure 2. Sequence for updating the user profile and user profile space.
a. The user starts a new application.
b. Manual adaptations made by the user are stored to the user profile.
c. Updating the user profile also updates the user profile space.

C. *User Interface Matching*

For the matching of a user interface based on a user profile, two different means of adaptation should be considered:

- Static configuration of (assistive) tools, which enhance an application (e.g., screenreader)
- Direct customization of an application (e.g., changing font size, or reordering menus)

For direct customization, an interface matching component obtains a list of possible adaptations of an application. From this list, it provides an optimal adaptation, as it compares the customization options with the user profile. The result is passed on to a special service (interface adaptor), which is responsible for performing the adjustments on the corresponding platform or application (see fig. 3).

For the matching itself, different approaches can be used: rule-based, statistical or a hybrid approach combining the rule-based and statistical approaches.

In the statistical approach, appropriate algorithms from the field of machine learning and statistical analysis will be used for the matching. A classical candidate approach is based on Support Vector Machines [9].

They perform a classification that can make a statement about which adaptation options make a good match to a user profile. This classification is based on previous adjustment decisions made by the user or similar users. Therefore, extensive training datasets must be created before this approach yields reliable results.

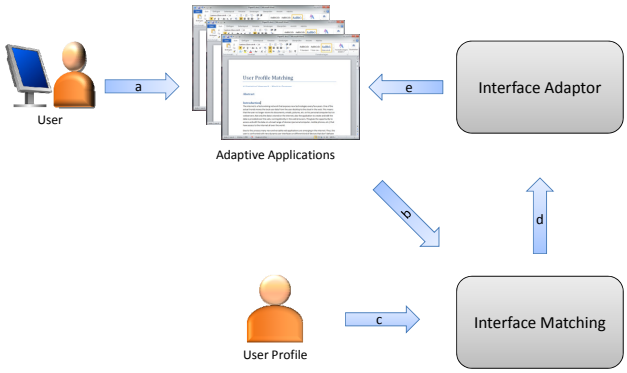


Figure 3. Sequence for matching a user interface to a user profile
a. The user starts a new application.
b. The application starts the user interface adaptation by transfering the adaptation feature list to the interface matching algorithm.
c. The matching algorithm also receives the user profile of the current user to perform the matchmaking.
d. The matchmaking results are forwarded to the interface adaptor.
e. The interface adaptor applies the adaptations on the running application.

In addition, the user can correct wrong decisions of the classifier by manually undoing or changing adjustments. These user adjustments are stored in the user profile and will influence future decisions of the matching algorithm.

IV. EVALUATION

The implementation and fine-tuning of the proposed approach requires an iterative approach based on real user data and extensive assessments of the results achieved.

At the beginning, we will develop an initial set of training and validation data. Based on this data, we will experiment with existing statistical methods and their parameters. The results of these approaches will be compared against the set of validation data. As more data becomes available, we will conduct additional optimization through parameter adjustments, and conduct small-scale user tests to compare this approach to the other approaches (rule/heuristics-based, hybrid, expert).

We will use manual adaptations by user interface experts as a benchmark to compare and evaluate the automatic adaptation strategies. The quantitative evaluation will be based on a compound metric, consisting of measurements on effectiveness, efficiency and user satisfaction. Our goal is for the statistical approach to come as close as 80% to the benchmark of the expert-provided adaptations.

V. OUTLOOK

In this paper, we introduced a concept for adapting user interfaces based on user profiles using statistical methods. For this, not only the matching algorithm is important, but also user profiles and their structure, collecting and storing user information and the classification of new user profiles within a user profile space.

 ISBN: 978-1-61208-167-0

One of the main tasks is to find optimal statistical methods for the various problems introduced in this paper. In order to do this, statistical approaches will be implemented on real profile data, applied, evaluated and compared among each other.

We intend to research and implement the described concept for matching between user interfaces and user profiles in a future European project. Together with other project partners we will also investigate additional approaches, including a rule-based and a hybrid approach, which combines the rule-based with the statistical approach.

REFERENCES

[1] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2007.

[2] T. Segaran, Programming Collective Intelligence, O'Reilly, 2007, pp. 7-28.

[3] T. Segaran, Programming Collective Intelligence, O'Reilly, 2007, pp. 197-224.

[4] Mainstreaming Accessibility through Synergistic User Modelling and Adaptability (MyUI). European FP7 project. http://www.myui.eu/

[5] Gentle User Interfaces for Elderly People (GUIDE). European FP7 project. http://www.guide-project.eu/

[6] A. Kobsa, Generic User Modeling Systems. User Modeling and User-Adapted Interaction 11, 2001, pp. 49-63.

[7] J. Sobecki, Consensus-Based Hybrid Adaptation of Web Systems User Interfaces. Journal of Universal Computer Science 11, no. 2, 2005, pp. 250-270.

[8] I. T. Jolliffe, Principal Component Analysis, 2nd ed., Springer, 2002.

[9] B. Schöllkopf, A. J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, The MIT Press, 2001.

 ISBN: 978-1-61208-167-0

Serious Games to Improve the Physical Health of the Elderly: A Categorization Scheme

Jaime A. Garcia Marin
mHealth Laboratory, iNEXT
University of Technology Sydney, FEIT
Sydney, Australia
Jaime.A.GarciaMarin@student.uts.edu.au

Karla Felix Navarro, Elaine Lawrence
mHealth Laboratory, iNEXT
University of Technology Sydney, FEIT
Sydney, Australia
Karla.FelixNavarro;Elaine.Lawrence@uts.edu.au

***Abstract*— this paper aims to provide a snapshot of the current status in the field of serious games for improving the physical health of the elderly. This work covers recent research projects for stroke rehabilitation and for falls prevention where user-center design methodologies were applied in order to satisfy this audience. A classification of the most relevant work in this area is provided along with a brief description of the platform, technology required and user-center design principles applied.**

Keywords- serious games, classification, categorization scheme, elderly, games for health, user-center design .

I. INTRODUCTION

Based on recent studies, it is stated that the aged population is dramatically increasing in both developing and developed countries. As an example, in Australia the aged population (+65 years old) increased by 94,800 people between June 2009 and Jun 2010, representing a 3.3% increase. In the last decade the aged cohort increased from 11.1% to 13.5%. These increases were also reflected in people over 85 years of age by 6.1 per cent and people over 100 years by 18.2 percent [1]. A government report predicts that the health spending will increase at the same rate as the number of aged Australians which is expected to double over the next four decades [2]. Therefore, the need of supporting this upcoming population has become a concern for governments and health providers around the world.

As a consequence of aging, the human body suffers a series of changes that could lead to the decline of the mental and motor capabilities. Furthermore, diseases such as postural instability, balance disorders and stroke are common at this age and are considered the main cause of disability among the elderly. [3, 4]

Current rehabilitation methods normally require the user to perform repetitive activities in order to recover lost motor abilities. However, the mechanics of this practice are often boring for patients affecting their motivation and commitment to the treatment [3].

Videogames have become popular among the aged population during the last decade, especially since the release of the Nintendo Wii [5]. Its revolutionary remote controller has changed the way to interact with videogames encouraging players to perform physical activities while playing [5].

Specialists and researchers have applied efforts to include interactive games in health treatments looking for a suitable method to keep their patients engaged. According to the literature, a significant number of studies have been conducted showing a positive impact among elderly users who increased their motivation and adherence to rehabilitation [6]. However, this practice could also result in undesirable consequences or poor outcomes for elder players when the aged-related changes are not considered [7, 8]. Usability issues have been found mainly because these games are not designed for this audience leading to negative impacts for the elderly [7-10].In view of that, most recent research projects are more concern about the end user needs, preferences and limitations during the design process as a strategy to guarantee accessibility [11-13]. In fact, the literature suggests that focusing on the intended user and their requirements has shown satisfactory results in games [14, 15].

As the main interest of this work is the usage of video games as an effective tool for improving the health of the elderly, one of the main goals of our research is to identify how the proper design could guarantee optimal results. The work presented here aims to provide a snapshot of the current status in the area trough a literature review and categorization of relevant work. Therefore, special attention was given to research projects that presented tested working prototypes and that incorporated user-centered design for the elderly.

The rest of the paper is structured as follows. Section 2 presents a brief summary of the evolution of this area including previous relevant classifications of this discipline. In Section 3, we describe our methodology and present our criteria of selection in Section 4. Section 5 contains a review of the games and finally, the discussion and conclusions are found in Section 6.

II. RELATED WORK

Over the last decade, the usage of videogames has become popular among different audiences. However such

 ISBN: 978-1-61208-167-0

video games have a huge potential as a tool for other purposes. "Serious games" is a relatively new field that could be described as the usage of videogames to help users to achieve a specific goal by playing a game [16]. This concept has been used to develop games for a range of areas such as such as medicine, defense, education and health among others.

In 2002, the 'Serious Games' initiative was formed to establish a formal basis for this emerging industry [17]. Two years later, the 'Games for Health' project was founded [18]. Its objective was to support community, knowledge and business development efforts to use games to improve health and health care. As a result of this movement, an annual conference is celebrated. The annual 'Games for Health' conference covers topics such as: exergaming, physical therapy, rehabilitation and training. In 2006, the ElderGAMES project started [19]. It was based on the use of entertainment in leisure time of the elderly as a tool for rehabilitation and prevention of the common diseases at this age. The project created an interactive-play board that aimed to maintain the elderly cognitive abilities through exercises [13].

In 2008, Sawyer and Smith [15] presented a categorization of serious games. This work establishes a snapshot of the state of this area at that time. The authors presented a general categorization of games and the application fields as well as further details for each category. Within the range of games for health, they present a classification of existing games grouped by intended audience and purpose.

In 2010, Rego et al. [16] extended this work presenting a taxonomy that focuses on games for rehabilitation. This survey covers a range of games to serve people with declines in motor and cognitive capabilities. They identified a set of criteria for the classification that focused on the following dimensions: (1) the purpose of the game, as the game could be intended for cognitive rehabilitation or physical rehabilitation; (2) the way the user interacts with the game; (3) either the game interface was two dimensional or three dimensional; (4) if the game allows more than one simultaneous players; (5) if the challenge was dynamically adapted based on the patient performance; (6) if the system enables the patient to know their progress; (7) progress monitoring capability; (8) capacity of game portability. At the end the authors compared the most relevant work found and used a system called RehaCom as a reference.

This work complements and extends previous classifications by presenting an expert categorization scheme where the use of games for improving the physical health of the elderly is the focal point. Consequently, the existing literature was surveyed focusing on games with this purpose. The inclusion of user centered design as a key element on this survey was incorporated as a result of previous studies conducted by the authors [20-22] as well as for its prevailing highlighted importance on existing literature in the area [11, 14, 15]. On account of the above the following topics were incorporated in this work: (1) games for post-stroke rehabilitation; (2) games for balance training and falls prevention; (3) user-centered design methodologies and evaluation; (4) guidelines for designing games for older adults. It is valuable to mention that stroke rehabilitation and fall prevention prevail through the literature , as stroke and falls are the most common causes of disability among the elderly [3, 23]. The next section set out the methodology used for our classification and the most relevant findings in terms of usability and gaming design for the elderly.

III. METHODOLOGY

A comprehensive search was conducted in order to gather relevant information for our categorization.

Searching for relevant data, peer reviewed journal articles (such as from ELSEVIER Health) and technical articles sourced from databases, for example, IEEE, ACM, EBSCO were reviewed for this work. The first phase of our search focused on identifying key points for a proper design for the elderly. That included understanding the ageing process and changes in the human body and common diseases at this age.

Based on our findings from stage one, we oriented the second phase of this work on technologies that included the entertainment factor to provide tools for physical and cognitive rehabilitation in patients that suffered stroke and games for falls prevention and balance training in older adults.

As our review was of a heterogeneous nature, a concept matrix was created in order to ease the classification of the existing games for improving the physical health of the elderly. For each reviewed project, the following information was registered:

1) *Audience:* As we are focused on games for health and the elderly, one of our goals was to identify if the game was suitable for older adults. This includes testing phases with older users and design evaluation post-playing.
2) *Goal:* this dimension is related to which area of the human body perceives benefits from playing the game, Also it aims to show if the game is designed as a training tool or for rehabilitation and if the main purpose was to either improve physical functions or cognitive processes.
3) *Interaction:* the way the user interacts with the game. Some systems use commercial input devices such as a remote controller, keyboard, mouse, etc. Some others developed their own input devices or ask the user to wear sensors that could be recognized by the platform.
4) *Technology:* a brief description of components and techniques that make it possible to run the game.
5) *Special Age Appropriate Features:* enhancement or mechanism that makes the game suitable for the elderly such as adapting the difficulty of the game dynamically.

 ISBN: 978-1-61208-167-0

6) *Home-based: (yes/no)* as many games are developed to be used at medical centers or rehabilitation centers. It is worth mentioning that games to be played at home could be more beneficial to users because this reduces the need to travel to rehabilitation centers.
7) *Feedback:* this dimension is related to the kind of feedback provided. Some games only use images and graphics on the screen to represent user actions (visual). Some others emit sounds (audio) or use vibration alerts (haptics) to notify actions to the user.
8) *Measurement:* this dimension is related to measurements that can be obtained during or after playing the game and could help specialists to determine improvements through medical assessment.
9) *Progress Record:* game characteristics that record user results (score, performance, etc).

IV. CRITERIA OF SELECTION

Overall, the reviewed literature showed that using games as a tool for rehabilitation and training has shown a positive outcome for the elderly. However, many tests have revealed that seniors could perceive this practice as an unpleasant experience and unfavorable results could be obtained. The main cause of this lies on usability issues related to changes in the human body over the years. Therefore, much effort has been applied to establish guidelines in order to create suitable games for this audience [24].

IJsselsteijn et al. [10] and presented a compilation of age related changes that must be considered when designing digital games for elderly users. They state that, although each individual differs from others in terms of abilities and experience, the human body normally tends to suffer a series of changes in sensory-perceptual processes, motor abilities and cognitive processes when getting older. For these reasons, the elderly user may not find these games enjoyable or beneficial if the games are not properly designed.

Flores et al. [9] conducted a search of journals and databases as a methodology to gather information regarding the most important game design principles for post-stroke rehabilitation. This review was focused on finding a set of criteria for both; a) designing effective therapy for post-stroke patients and b) entertainment for the elderly. It was found that the most of the games for post-stroke rehabilitation did not include enjoyable content for the elderly. Furthermore, they proposed a set of criteria for both which is set out in Table 1.

Based on these guidelines, key points for suitable design for the elderly were identified and used for classification. Thus, the concept matrix was used as a tool for filtering data and identifying the most relevant work. Our criteria of selection are set in Table 2.

Table 1: Gaming design criteria for stroke rehabilitation programs serving elderly users [9]

Criteria for Stroke Rehabilitation	Criteria for Elderly Entertainment
• Adaptability to motor skill level • Meaningful tasks • Appropriate feedback • Therapy-Appropriate Range of Motion • Focus diverted from exercise	• Appropriate cognitive challenge • Simple objective/interface • Motivational Feedback • Element of social activity • Appropriateness of genre • Creation of new learning following guidelines of experts • Sensitivity to decreased sensory acuity and slower responses

Table 2: Criteria of Selection

Parameter	Value
Audience	Elderly
Goal	Improve physical health (upper limbs, lower limbs, balance)
Interaction	Any of these: Shifting weight, Wearing Sensor (Image Recognition), Stepping on Surfaces, Touching Surfaces, and Grasping Objects.
Technology	Any of these: commercial platforms (Wii, PS, Xbox, Kinect), PC Games, Robot, Balance Board, commercial remote controller, Camera / WebCam, Dancing Pad, MultiTouch Tabletop.
Special Age Appropriate Features	Large visual instructions, audio assistance, mechanisms to dynamically adapt challenge, monitored by the Occupation Therapist.
Home-based	Preferable
Feedback	Any of these: Visual, Haptics, Audio
Measurement	Any of these: range of motion, user movements and trajectories, high scores, game results.
Progress Record	Preferable

V. REVIEW OF GAMES FOR IMPROVING THE PHYSICAL HEALTH OF THE ELDERLY

This section focuses on providing a brief description of each game(s), the range of technology used in this field, interactions with the games and usability studies and relevant results. As it was mentioned before, much attention was given to working prototypes that followed a user-centered design methodology. We present here some of the research projects that meet the criteria of selection. They are grouped by as follows: (1) games for improving lower-limbs functions and balance training (see Table 3). (2) Games for rehabilitation / exercising of upper-limbs (see Table 4).

 ISBN: 978-1-61208-167-0

Table 3: Games for balance training and for lower-limbs rehabilitation / exercising.

	Game	Interaction	Technology	Special Feature	home-based	Feedback	Measure	Progress Record
Smith et al [25]	DDR	Stepping on pad	PC Dancing Pad	Music Thematic of the game Share progress with OT (Number of Participants not stated)	Yes	Visual Audio	Game Results	Yes
Doyle et al. [26]	Otago Exercises	Shifting weight	PC Wii Balance Board	Monitored by therapist Can skip non-playable content (6 Participants)	Yes	Visual Audio	Game Results	Yes
Gerling et al [27]	Task 1	Shifting weight	PC Wii Balance Board	Simplistic design single tasks adaptive difficulty (9 Participants, Ages 77-91)	Yes	Visual	Game Results High Scores	Yes
	Task 2	Jumping						

Smith et al. [25] uses a modified version of the game Dance Dance Revolution. The main purpose of this game is to provide a tool to train the stepping abilities of the elderly, a common problem experienced by this population. In order to interact with the game, the player must step on a dance pad sensor that has eight arrows. A display (TV or PC Monitor) provides step direction instructions to the player by scrolling arrows from the bottom to the top of the screen. The game is adapted for slower responses the elderly. Also, the author presents a design for monitoring the user performance by using mobile technology. This aims to enhance cooperation among patients and therapists by sending information about user performance to the practitioner. Finally, the system is designed to be used at home.

Doyle et al. [26] developed a game to deliver balance and strength exercises. This project aims to help elderly users to improve their motor capabilities in lower-limbs in order to avoid falls. This system is made of a flash application running on a Laptop, a camera (webcam), a set of markers for upper and lower limbs and SHIMMER kinematic sensor for walking exercises. The game provides five exercises from the OTAGO exercise programme [28]. As the user performs the exercises, his/her performance is remotely monitored by the instructor in order to validate the correct completion of each exercise. Additionally, the author conducted a series of usability tests in order to identify the user preferences to make the platform more attractive to elderly users. These tests focused on evaluating visual and audio feedback, navigating through the application, providing instructions and measuring attitudes / motivation to exercise. The most important aspect found were: (1) users are more likely to play games that use an avatar instead of seeing themselves on screen (visual feedback). (2) With audio feedback, users prefer counters that emit single sounds (like: 'ding') instead of listening to number countdowns. The latter could be distracting for patients leading to lack of concentration. (3) It was found that providing options to pause / resume the game or skip to the next exercise eases the navigation through the application. (4) The notion of being under observation by a therapist increases the motivation of the patient.

Gerling et al. [27] present a case study where they developed a game for balance training considering the needs of the elderly. Their prototype, called SilverBalance, uses the Wii Balance Board and consists of two single tasks with a simplistic graphic design. In task 1, a series of obstacles randomly appear aligned to the left or right and the user must shift weight to the opposite way in order to avoid the obstacles. As long as the user plays, the speed increases until the player is not able to achieve the goal. In task 2, the obstacles cover the width of the screen and the user must 'jump' to avoid collisions. At the end of each task, the system shows the user performance and saves high scores. Both activities can be performed either sitting or standing, which is more accessible for this audience and allows people in wheelchairs to participate. The focus group test was composed of nine older adults with an average age of 84. After testing the usability of this prototype, it was found that simplistic designs allow the user to concentrate on the game encouraging them to perform the exercises. Also, this work shows that applying design principles has a positive outcome for elder users.

 ISBN: 978-1-61208-167-0

Table 4: Games for upper-limbs exercising / rehabilitation.

	Game	Interaction	Technology	Special Feature	home-based	Feedback	Measure	Progress Record
Burke et al. [3]	RabbitChase Bubble Trouble ArrowAttack	Hand movements wearing gloves (image recognition)	PC Camera Webcam Markers	Mechanism to adapt difficulty based on user performance Tool for analyzing recorded log files (3 Participants, Ages 65-73)	Yes	Visual Audio	Player's movement trajectories Range of motion,	Yes
Fasola & Mataric [29]	Workout game Imitation game Memory games	Image Recognition without Markers	Robot Camera Wiimote	Social Factor Motivation (11 Participants, Ages > 65)	No	Audio Visual	Hand location, Arm angles	Not mentioned
Annet et al.[30]	Pop those balloons Drumhab Paint by number Picture tracing Therapist Do-It-Yourself	Touching screen	PC MultiTouch tabletop	Therapist can modify the difficulty of the game, changed touch sensitivity (Number of participants not stated)	No	Visual	Touch pressure	Not mentioned

Burke et al. [3, 31] developed a series of webcam games considering the theory for design and rehabilitation (meaningful games, appropriate challenge). This project aims to provide a low-cost tool for upper limbs rehabilitation that can be used at home. This platform was developed using a commercial development kit and libraries and requires a webcam, a PC and colored gloves in order to operate. Four games are provided:

1) RabbitChase: the game presents four holes and one rabbit. Eventually, the rabbit comes out of one hole, walks and gets in another hole, both randomly chosen by the game. The player must point at the hole into which the rabbit hid. If the player is correct, encouraging visual/audio feedback is given. Also, these researchers developed a mechanism that automatically adapts the difficulty of the game based on patient success.
2) Bubble Trouble: Floating bubbles randomly appear on the screen, then after a short period of time they disappear. They user must touch them before they disappear, making them burst.
3) ArrowAttack: This game shows two arrows that are colored according to the user gloves or markers. One points to the left and the other to the right. Also four boxes are shown. The arrows move from one box to another and the user must imitate this movement with their hands as long as they move.

The system also includes a tool that analyses saved log files that are given to the therapist. Additionally, two playability studies were conducted. It was found that the adaptive mechanism that increases and decreases the difficulty of the game, was too aggressive when adjusting the challenges so it must be refined to be more 'gentle'.

Fasola & Mataríc [29] implement an assistive robot to deliver arm exercises for the elderly. This robot monitors the performance of the user and provides motivation to the player promoting an increased range of motion. In order to operate, the user must sit in front of the robot and three different games are given:

1) *Workout game*: the robot acts as a traditional instructor giving a series of exercise that the user must perform;
2) *Imitation game*: the player acts as the instructor and the robot imitates the user movements; and
3) *Memory games*: the robot provides a sequence of arm gestures and the player must memorize them and repeat them.

In order to capture the user movements, the research team developed a vision module that recognizes the users' faces and determines hand location. In order to simplify the visual recognition, they installed a black curtain behind the user to provide a contrasting and static background. At the end of the trials, a survey was conducted to determine the participants' feeling and perceptions towards the robot. The results suggest that

 ISBN: 978-1-61208-167-0

the participants perceive the robot as a trustable entity able to help them in exercising training.

Annet et al. [30] developed a multi-touch table system to deliver training for upper-limbs. This platform was built under the guidance of occupational therapists, specialists who normally work with patients to help restore or improve motor functions. Three objectives were established at the beginning: (1) engage the user and provide easy to learn activities; (2) create repeatable activities, measure user performance and record it; (3) build on the therapist's expertise and knowledge of a patient. The whole system is made of an existing multi-touch technology in conjunction with a set of applications for rehabilitation. Five games are given:

1) *Pop those balloons*: a landscape and a number of floating balloons are shown. The user must touch them in order to pop the balloons. Once the player touches the balloon, it disappears and increases the score. The therapist is able to modify the number of balloons and their speed while the patient is playing the game, in order to increase or decrease the difficulty of the activity.
2) *Drumhab*: this game presents a center orb that emits 'beats' from the middle of the screen to the four corners. Each corner has a drum that must be touched when a beat reaches its position. If the player hits the drum at the correct time, the 'beat' disappears. The therapists can control the game by changing the number of beats and their speed as well as which drums are targets.
3) *Paint by number*: this game shows a group of paint buckets and an outline with numbers. Each paint bucket has a number and a different color. The user must use their hand as a paintbrush and paint the image.
4) Picture tracing: In this game, the therapist draws a pattern on the surface of the tabletop and the patient is asked to trace overtop of the pattern.
5) Therapist Do-It-Yourself: the therapist creates a sequence of points (targets) and the user must reach them. Once he/she touches a target, it disappears and the next one comes up.

VI. DISCUSSION AND RESULTS

One of the main limitations for this work was the lack of research across the area of the elderly and the user of serious games for health purposes. For that reason, literature from intersecting disciplines such as the use of games for stroke treatment and games for improving balance function was reviewed to obtain more data. In spite of that, it is valuable to mention that some work was conducted in the late eighties and early nineties. Some improvements in cognitive processes were observed after conducting trials with elderly patients who played specifically selected computer games [32, 33].

On our previous research on the Nintendo Wii, the suitability and usability of six commercial balance games were assessed by three conventional health professionals and three alternative health professionals [11, 14, 15]. They identified a series of problems that could lead to negative results for the elderly cohort, as their needs were not considered during the design process. For instance, the health professionals stated that the elderly could learn inappropriate movements in response to game actions; resulting in risks for their physical condition. Also, they stated that hearing impairments and sight problems were not properly addressed when providing rules and instructions which could lead to confusion and frustration. However, we have learnt that current rehabilitation techniques that apply user-centered design principles are being well accepted by older patients, especially when the games and exercises involve fun factors. This highlights the importance of adequate design for this audience in order to obtain positive results.

Additionally, we noticed that much effort is applied on developing games for improving upper-limb function. Some of the noted reasons for that in the literature were: (1) daily activities require the use of arms and hands such as grasping objects and brushing teeth among other daily activities [4, 34]. (2) A significant number of elder users could require a wheelchair, so they would not be able to perform activities that require the whole body [27, 35, 36].

In addition to the above, it is important to take into account that the use of webcams has been a popular low-cost input device solution. The nature of the device imposes inherent technological limitations to capture accurately the whole body. It is worth mentioning that costing of those games was not mentioned as the researchers developed extra items such as robots or mechanical arms and employed programmers to assist in the games' construction.

Regarding the usage of commercial platforms for this purpose, we found that there is a tendency towards the utilization of their input devices rather than the whole system. Although these innovative ways to interact increase the accessibility and motivation of patients, the coverage could be limited due to the lack of standardization among platforms.

Finally, it was observed that even though most of the studied prototypes included progress monitoring and some type of measurement [3, 25], only a small number of these leverage on the therapist knowledge to assess the effectiveness of the therapy [30]. Hence, future work on the inclusion of mechanisms to perform medical validation and assessment in real-time could bring important benefits for patients and therapists alike.

 ISBN: 978-1-61208-167-0

VII. CONCLUSIONS

In this paper we have classified the work of serious games for the elderly according to their specific goal and key design elements for improving the physical health of the target audience. This expert categorization scheme aims to help researchers in the area to identify the current status and major needs in this new field of research. It was found that current projects tended to have an increased awareness of the elder user needs in gaming design for health purposes. Yet, little effort has been made to take into account the limitations of the aged cohort along the entire design process. This could be attributed to the favorable results that have been obtained in the past on the incorporation of elderly specific human centered design guidelines to assists the design process. However, its theoretical nature could not be sufficient to guarantee optimal results, so testing phases and usability evaluation are needed. As a final point, the use of modern input devices has allowed specialists to determine the patient performance in terms of motor functions. Nevertheless, the potential of use this capabilities to accurately perform medical assessment as a tool to guarantee effectiveness is not being exploited. This has been identified by the authors as an important direction for future research in the area and it is intended to be the next stage of our research.

VIII. REFERENCES

[1] (2010, May 7, 2011). *Population by Age and Sex, Australian States and Territories, Jun 2010* Available: http://www.abs.gov.au/Ausstats/abs@.nsf/mf/3201.0

[2] E. Rodgers. (2010, May 17, 2011). *Ageing population dragging Australia into the red*. Available: http://www.abc.net.au/news/stories/2010/02/01/2806592.htm

[3] J. Burke, M. McNeill, D. Charles, P. Morrow, J. Crosbie, and S. McDonough, "Optimising engagement for stroke rehabilitation using serious games," *The Visual Computer,* vol. 25, pp. 1085-1099, 2009.

[4] B. Cesqui, S. Micera, S. Mazzoleni, M. C. Carrozza, and P. Dario, "Analysis of Upper Limb Performance of Elderly People Using a Mechatronic System," in *Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. The First IEEE/RAS-EMBS International Conference on*, 2006, pp. 365-370.

[5] Y.-L. Theng, P. Teo, and P. Truc, "Investigating Sociability and Affective Responses of Elderly Users through Digitally-Mediated Exercises: A Case of the Nintendo Wii," in *Human-Computer Interaction*. vol. 332, P. Forbrig*, et al.*, Eds., ed: Springer Boston, 2010, pp. 152-162.

[6] H. Sugarman, A. Weisel-Eichler, A. Burstin, and R. Brown, "Use of the Wii Fit system for the treatment of balance problems in the elderly: A feasibility study," in *Virtual Rehabilitation International Conference, 2009*, 2009, pp. 111-116.

[7] S. Hanneton and A. Varenne, "Coaching the Wii," in *Haptic Audio visual Environments and Games, 2009. HAVE 2009. IEEE International Workshop on*, 2009, pp. 54-57.

[8] C. Neufeldt, "Wii play with elderly people," IISI2009.

[9] E. Flores, G. Tobon, E. Cavallaro, F. I. Cavallaro, J. C. Perry, and T. Keller, "Improving patient motivation in game development for motor deficit rehabilitation," presented at the Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, Yokohama, Japan, 2008.

[10] W. Ijsselsteijn, H. H. Nap, Y. d. Kort, and K. Poels, "Digital game design for elderly users," presented at the Proceedings of the 2007 conference on Future Play, Toronto, Canada, 2007.

[11] E. Kalapanidas, C. Davarakis, F. F. Aranda, S. Jiménez-Murcia, O. Kocsis, T. Ganchev, H. Kaufmann, T. Lam, and D. Konstantas:, "PlayMancer: Games for Health with Accessibility in Mind," *Communications&Strategies,* vol. 73, pp. 105-120, 2009.

[12] L. Gamberini, M. Alcaniz, G. Barresi, M. Fabregat, F. Ibanez, and L. Prontu, "Cognition, technology and games for the elderly: An introduction to ELDERGAMES Project," *PsychNology Journal,* vol. 4, pp. 285-308, 2006.

[13] L. Gamberini, F. Martino, B. Seraglia, A. Spagnolli, M. Fabregat, F. Ibanez, M. Alcaniz, and J. M. Andres, "Eldergames project: An innovative mixed reality table-top solution to preserve cognitive functions in elderly people," in *Human System Interactions, 2009. HSI '09. 2nd Conference on*, 2009, pp. 164-169.

[14] A. Väätänen and J. Leikas, "Human-Centred Design and Exercise Games," in *Design and Use of Serious Games*. vol. 37, M. Kankaanranta and P. Neittaanmäki, Eds., ed: Springer Netherlands, 2009, pp. 33-47.

[15] V. A. V. Abeele and V. V. Rompaey, "Introducing human-centered research to game design: designing game concepts for and with senior citizens," presented at the CHI '06 extended abstracts on Human factors in computing systems, Montreal, Quebec, Canada, 2006.

[16] P. Rego, P. M. Moreira, and L. P. Reis, "Serious games for rehabilitation: A survey and a classification towards a taxonomy," in *Information Systems and Technologies (CISTI), 2010 5th Iberian Conference on*, 2010, pp. 1-6.

[17] (May 17, 2011). *Serious Games Initiative*. Available: http://www.seriousgames.org/

[18] (2004, May 16, 2011). *Games for Health Project*. Available: http://www.gamesforhealth.org/

[19] (2006, May 16, 2011). *The ElderGames Project*. Available: http://www.eldergames.org/

[20] K. M. F. Navarro, E. M. Lawrence, J. G. Marin, and C. Sax, "A Dynamic and Customisable Layered Serious Game Design Framework for Improving the Physical and Mental Health of the Aged and the Infirm," presented at the Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED 2011), Gosier, Guadeloupe, France, 2011.

[21] J. G. Marin, E. M. Lawrence, K. M. F. Navarro, and C. Sax, "Heuristic Evaluation for Interactive Games within Elderly Users," presented at the Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED 2011), Gosier, Guadeloupe, France, 2011.

[22] E. M. Lawrence, K. M. F. Navarro, J. G. Marin, and C. Sax, "Towards building health systems," presented at the The Sixth International Conference on Systems (ICONS 2011), St. Maarten, The Netherlands Antilees, 2011.

[23] D. Lloyd-Jones, R. Adams, T. B. M. C. M, S. Dai, G. D. Simone, T. Ferguson, E. Ford, K. Furie, C. Gillespie, A. Go, K. Greenlund, N. Haase, S. Hailpern, P. Ho, V. Howard, B. Kissela, S. Kittner, D. Lackland, L. Lisabeth, A. Marelli, M. McDermott, J. Meigs, D. Mozaffarian, M. Mussolino, G. Nichol, V. Roger, W. Rosamond, R. Sacco, P. Sorlie, V. Roger, T. Thom, S. Wasserthiel-Smoller, N. Wong, and J. Wylie-Rosett, "Heart Disease and Stroke Statistics - 2010 Update. A Report From the American Heart Association," *Circulation,* vol. 121, Feb 23, 2010 2010.

[24] L. Gamberini, M. Alcaniz, G. Barresi, M. Fabregat, F. Ibanez, and L. Prontu, "{Cognition, technology and games for the elderly: An introduction to ELDERGAMES Project}," *PsychNology Journal,* vol. 4, pp. 285-308, 2006.

[25] S. T. Smith, A. Talaei-Khoei, M. Ray, and P. Ray, "Electronic Games for Aged Care and Rehabilitation," in *e-Health*

 ISBN: 978-1-61208-167-0

Networking, Applications and Services, 2009. Healthcom 2009. 11th International Conference on, 2009, pp. 42-47.

[26] J. Doyle, C. Bailey, B. Dromey, and C. N. Scanaill, "BASE - An interactive technology solution to deliver balance and strength exercises to older adults," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on-NO PERMISSIONS*, 2010, pp. 1-5.

[27] K. M. Gerling, J. Schild, and M. Masuch, "Exergame design for elderly users: the case study of SilverBalance," presented at the Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology, Taipei, Taiwan, 2010.

[28] N. Z. s. A. C. Corporation. (2009). *Otago Exercise Programme*. Available: http://www.acc.co.nz/preventing-injuries/at-home/older-people/information-for-older-people/otago-exercise-programme/index.htm

[29] J. Fasola and M. J. Mataric, "Robot exercise instructor: A socially assistive robot system to monitor and encourage physical exercise for the elderly," in *RO-MAN, 2010 IEEE*, 2010, pp. 416-421.

[30] M. Annett, F. Anderson, D. Goertzen, J. Halton, Q. Ranson, W. F. Bischof, and P. Boulanger, "Using a multi-touch tabletop for upper extremity motor rehabilitation," presented at the Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7, Melbourne, Australia, 2009.

[31] J. W. Burke, M. D. J. McNeill, D. K. Charles, P. J. Morrow, J. H. Crosbie, and S. M. McDonough, "Serious Games for Upper Limb Rehabilitation Following Stroke," in *Games and Virtual Worlds for Serious Applications, 2009. VS-GAMES '09. Conference in*, 2009, pp. 103-110.

[32] G. R. Whitcomb, "Computer games for the elderly," *SIGCAS Comput. Soc.,* vol. 20, pp. 112-115, 1990.

[33] J. Goldstein, L. Cajko, M. Oosterbroek, M. Michielsen, O. Van Houten, and F. Salverda, "VIDEO GAMES AND THE ELDERLY," *Social Behavior & Personality: An International Journal,* vol. 25, p. 345, 1997.

[34] J. P. Giuffrida, A. Lerner, R. Steiner, and J. Daly, "Upper-Extremity Stroke Therapy Task Discrimination Using Motion Sensors and Electromyography," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 16, pp. 82-90, 2008.

[35] B. Herbelin, J. Ciger, and A. L. Brooks, "Customization of gaming technology and prototyping of rehabilitation applications," presented at the International Conference Series On Disability, Virtual Reality And Associated Technologies, Portugal, 2008.

[36] H. Zabaleta, M. Bureau, G. Eizmendi, E. Olaiz, J. Medina, and M. Perez, "Exoskeleton design for functional rehabilitation in patients with neurological disorders and stroke," in *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, 2007, pp. 112-118.

 ISBN: 978-1-61208-167-0

CodeDroid: A Framework to Develop Context-Aware Applications

Lillian B. R. de Oliveira Antonio A. F. Loureiro
Department of Computer Science
Federal University of Minas Gerais
31270-010 Belo Horizonte, MG, Brazil
{*lillys,loureiro*}*@dcc.ufmg.br*

***Abstract*—Context-aware computing enables the development of systems that adapt themselves to the context of the user, device and surrounding environment. To represent the context information associated with the application, this work proposes the use of a profile. This paper also proposes the CodeDroid framework, based on the Android platform, to help a designer in the development of context-aware mobile applications for different domains. This strategy showed to provide greater modularity and reuse of components. The CodeDroid framework also brings together the most generic services used to collect context, as discussed in the development of Places2Go, a tourism application.**

***Keywords*-context-aware computing; interface; mobile application;**

I. INTRODUCTION

The use of mobile devices is a common aspect of our lives. These devices have more features, more processing power and communication capacity. With the increased use of mobile devices new computing platforms were designed for that environment. Among them we can mention the Android platform that deservers special attention since is becoming the dominant operating system in mobile devices. Due to the success of these platforms, there is an increase in the design of context-aware applications that make use of existing information collected from sensing devices and other user-related information. To facilitate the development of applications and achieve greater productivity, it is recommended to use *frameworks* or other similar computational tools that help designers to optimize their work. For example, a single application for a mobile computing environment may have to consider dozens of portable devices with different characteristics and it can be customized to meet the requirements of different users.

Another incentive for the development of mobile applications is related to the financial aspect. Developers can sell their own applications to users around the world through virtual markets available, for example, by manufacturers. This scenario shows how important the agility of developing mobile applications is, especially, in a competitive market such as mobile computing.

Context-aware computing allows the development of systems that adapt to the context. The context may include information from the user, device and environment. Context-aware applications rely on context to provide services to users, however, not all context information is relevant to the application. In this work, we employ the concept of profile to centralize information about relevant user profile data, as described in the next section.

Currently, the development of applications for the Android platform is not a simple task. The developer must understand the architecture, the execution flow of applications and characteristics of the platform components. For example, the code of an application on Android can be developed into a single component opposing the practice of internal cohesion and low coupling. To facilitate the use of the profile and organizing the architecture of context-aware applications we propose the CodeDroid framework that allows the development of context-aware mobile applications for the Android platform.

In the literature, there are some proposals for framework for context-aware applications developed on the Android platform [2], [3]. But they do not use the concept of centralized profile as information on the context. Other proposals do not address sensing services implementations used by the generic context-aware applications.

This work aims to use the concept of profile in the design of context-aware applications. We have developed the CodeDroid framework and applied it to design an application called *Places2Go* for a tourism case study. The application was evaluated using real and emulated environments with good results demonstrating the feasibility of the proposed strategy.

This work is organized as follows. Section II discusses the concept of profile used in this work. Section III gives a brief overview of the Android platform, which was used to develop our proposed solution. Section IV presents the CodeDroid framework that was designed to help designers in the development of context-aware mobile applications for the Android platform. Section V evaluates the proposed framework in both real and simulated environments. Finally, Section VI presents our conclusion and future work.

II. PROFILE

Dey et al. [1] define context as "any information that can be used to characterize the situation of entities (i.e., a person, place or object) that are considered relevant to the interaction between a user and an application, including

 ISBN: 978-1-61208-167-0

user and the application. Context is typically the location, identity and status of people, groups and computational and physical objects". To aggregate the information relevant to context, applications can use a concept called profile in this work.

Definition (Profile) *Entity responsible for representing the context information of a user, device and environment that is relevant to the interaction with the application*.

The profile is very important for context-aware applications because it helps to decide the flow of the application execution. The application starts interacting with the profile, so the the user and sensing services do not need to be frequently involved in these interactions. The profile centralizes for information and acts as interface between application and other services.

The contents of the profile can be modified at any time and this information is fundamental to offer for more adequate services to users. Such changes can be observed in situations where the actual user moves from one location to another one or when other users and/or resources move into and/or outside the application area of interest. The dynamics of the profile makes its use more complex, however, its constant update allows to better meet the goals of both user and application.

The profile structure is the way used to express the context information and allow its exchange. In the current scenario, each application uses a particular way of structuring and modeling context, which can be completely different from another application, even if they belong to the same domain. In this way, it becomes very complicated to exchange information among those applications. Thus, it is important to standardize the context representation to make easy the communication between user and application and among applications.

The structures to represent a profile can be divided into the following categories: model marking scheme, model of key-value pairs, object-oriented model, logic-based model and ontology. This work uses the model marking scheme that allows to represent information in a robust way, without demanding too much processing to treat the information. Another advantage is that the profile defined in this way allows to reuse it in different domains, what does not happen if we use ontologies, for example.

In the literature, there are several ways to represent context such as XML (*Extensible Markup Language*), RDF (*Resource Description Framework*), RDF-Schema and OWL (*Web Ontology Language*). Due to the choice of the markup model as a way of structuring the profile, we chose XML, which is widely used for information exchange, especially in web services. XML also works as the basis for other forms of representation mentioned above. Therefore, the designer can still use other representation forms provided they are described in XML.

The profile contents incorporates information about the user's context, device and environment. The user information may describe personal characteristics (e.g., name, age, sex and education) and interests. The device context includes the device specification (e.g., model, processor, display size, networking capabilities and existing sensors) and conditions (e.g., battery level, available memory and connected network). Some of the environment characteristics that may be present in the profile are weather conditions, noise levels and traffic conditions.

This paper addresses the design of an entity called "profile" with some predefined attributes. Since the framework goal is to support different domains, we chose the more generic attributes that can be used in various context-aware applications. These attributes were chosen based on the evaluation of various context-aware applications in different scenarios. The attributes considered to be the most generic to comprise the profile are: identifier, name, email address, location and battery level. However, the designer can also define additional attributes for the profile. This situation only occurs when the application needs to define more context data beyond those already listed in the profile of the framework. In this case, the designer is responsible for defining the attribute and collecting context data to assign to the attribute information.

To build the profile we have to collect data about the context. The data collection can happen in different ways: interactively, import from an existing data source or service sensing. The interactive form depends on the user who must provide the necessary data directly. In this case, the application may stay in an idle state waiting for a user response, hindering its operation. The form of collecting by importing data implies that the profile represents data from a file to be imported by the application. This import does not guarantee that the data is updated and requires a standard format for the context representation file.

Data collection through sensing services is highly recommended. It allows to obtain real and updated data. Another feature is that most of the devices already have the necessary sensors to provide services to access the hardware. We evaluated context-aware applications in different scenarios and identified the need to use the following basic services:
– *Location Service*: obtains the user's location. The resulting information is expressed as coordinates (latitude and longitude) or a full address.
– *Weather service*: Provides weather information about a given location.
– *Energy management service*: checks the availability of the battery energy.

These services can work independently from other services. However, some of them depend on other services. For example, the weather service depends on the information provided by the location service. The framework addresses this issue by allowing a dependency hierarchy defined by

 ISBN: 978-1-61208-167-0

the designer.

Another important aspect is the frequency and service execution time. It is possible to configure each service to perform in a given period of time, saving device's resources. It is the developer's responsibility to check the need to use a service and the consumption of battery.

III. Android Platform

We used the Android platform[1] to implement the CodeDroid framework. Among the advantages of this platform we highlight the portability, because it uses its own virtual machine (*Dalvik*). It is based on open code, which allows greater flexibility in the application development. It also has a robust development environment with tools for code debugging and performance evaluation. It also provides facilities to access device features (e.g., GPS, camera, battery, network, among others). To build applications for the Android platform, we use the following components:

Activity: most activities interact with the user. Therefore, the activity is responsible for creating the window in which the graphical user interface will be built. An activity may be presented as a floating window or embedded in another activity.

Service: recommended for functionalities that need to run indefinitely in the background. It has no visual interface.

Broadcast receivers: receives and responds to notifications sent to the system.

Content providers: component responsible for storing and retrieving data and making them accessible to all applications.

In the Android architecture there is no structure and organization that determines the communication pattern between components. All component types can instantiate components of the same type or not. For example, an activity can communicate with a service, a broadcast receiver or a content provider. It is possible to create Android applications using only one activity that connects to pre-existing services in the platform. In this case all code may belong to just one class, making it less cohesive, high coupling, less readable and more complex to maintain. Making an analogy between the Android architecture and the MVC architecture model, we can say that *Activity* performs the role of the vision and control layer. The *ContentProvider* would be the model layer.

IV. CodeDroid Framework

The goal of the CodeDroid framework is to assist the development of mobile context-aware applications for the Android platform. In addition to the contributions of a *framework* (extensibility, control inversion, modularity and reuse), CodeDroid allows a user to add other features through the standardization of the application structure: better organization of the application code; adaptation of the MVC pattern to develop Android applications, greater modularity; definition of an access point to the sensing services based on the *Facade* standard; implementation of notifications as a result of context changes based on the *Observer* design pattern; definition of the "entity" for representing context; profile abstraction allowing its management in a transparent manner; and providing services for the collection of context sensing.

A. Framework Architecture

The framework architecture is depicted in Figure 1. The classes of the CodeDroid framework are detailed below.

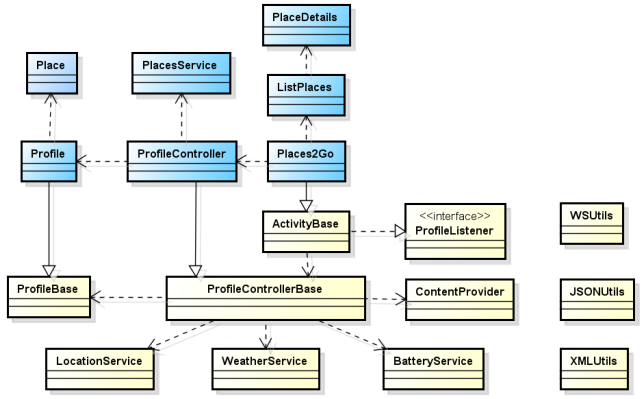


Figure 1. Class diagram of the application *Places2Go* and its relationship with the classes of CodeDroid framework.

ActivityBase: extends the class *Activity* already defined in Android. The goal is to relate all the basic settings of an *Activity* in a single class, which depends on the class *ProfileControllerBase* to use the other framework services. The application developer must define a main *Activity* for the application that extends the *ActivityBase* class. Thus, the developer can use the profile management service without knowing its configuration and take advantage of the framework.

ProfileControllerBase: aims to bring together all profile management services. It acts as the application control that receives the instructions of the *ActivityBase* class, and manages which operations will be performed. This class is also connected to the layer model that represents the profile.

The class *ProfileControlerBase* is the extension of *Service* component defined by Android. As the features already listed about this component, the class *ProfileControlerBase* can run even in the background. This class already implements the most used services in context-aware applications. The available services in the framework are: location service (*locationservices*), battery management service (*BatteryService*) and weather service (*WeatherService*).

When using the framework, the developer must create a subclass of *ProfileControllerBase*, which represents the

[1] http://www.android.com/

 ISBN: 978-1-61208-167-0

application control layer. It is important to notice that, besides the services available at the CodeDroid framework, the developer can add other services. The developer should preferably define each service in a class, ensuring a greater cohesion and a lower coupling. Additional services should be connected to the subclass of *ProfileControllerBase*, following the logic of the other services. This shows the flexibility and extensibility of the framework.

ProfileBase: default class that represents the entity profile. This class lists the essential attributes of the profile, i.e., those that are commonly used in context-aware applications in different scenarios. The class has the attributes: identifier, name, email address, location, and battery level.

ProfileListener: class that aims to manage the change notifications of the profile. The class *ActivityBase* implements this interface. Therefore, the main *Activity* of the application, which extends the class *ActivityBase*, must implement the methods defined in the interface. These methods will be responsible for processing events generated when the profile is modified.

For each event that occurs in pre-service implemented by the framework, the method *onProfileChanged* of the class *ProfileListener* is invoked. The developer should handle the events in the implementation of the method *onProfileChanged* of the main *Activity* application. Without this class, the developer would have to implement the listeners of all services needed by the application.

BatteryService: manages the battery level of the user's device. It captures the real information directly from the user's device. This is a subclass of the *Service* class available at Android, which runs in background. The class *ProfileControllerBase* manages this service and other sensing services as well.

LocationService: service responsible for obtaining location information from the user's GPS device. It is a subclass of the *Service* class.

WeatherService: returns the weather conditions of a particular region. This region may be the user's current location, obtained from the location service, or any other location such as the latitude and longitude coordinators or address. It is a subclass of *Service*.

WSUtils: class responsible for mediating the interaction between the CodeDroid framework and Web services, i.e., this class helps to integrate queries to Web services.

XMLUtils: class responsible for performing the processing of XML data. It interprets data in XML format that comes from the Web services and translates it to a format understood by the CodeDroid framework.

JSONUtils: class responsible for processing data in JSON format (*JavaScript Object Notation*). It interprets data in JSON format that comes from the Web services and translates it to a format understood by the CodeDroid framework.

B. Design Patterns

The framework CodeDroid has been developed based on design patterns. The standards used in architectural modeling of Android were MVC (*Model-View-Controller*), *Facade*, *Singleton* and *Observer*. The MVC architectural pattern was the starting point in the design of the CodeDroid framework. However, we found that the Android platform presents some shortcomings to use this model. The purpose of the CodeDroid framework is to organize the architecture of Android applications to follow the MVC model. The class *ActivityBase*, its subclasses and other activities represent the view layer. Another important component that is included in the view layer is the layout — XML files built from predefined tags that represent the visual components of the Android platform. The offered service by the class *ProfileControllerBase* together with their subclasses is the application control layer. The model layer is represented by the entity profile.

The class *ProfileControllerBase* was defined based on the standard *Facade*. It is responsible for intermediating the communication between clients and other services. That is, this class receives commands from an *Activity* and triggers the actions to the corresponding service.

The service *ProfileControllerBase* was also built based on the standard *Singleton*. We identified the need to define a unique service to perform profile management. The applications call this service whenever they need context information. All context-aware applications developed with the CodeDroid framework access the same service *ProfileControllerBase*. A single service ensures a better performance because it would be very difficult if, for each execution, the application would create different new services. Another advantage is that applications can take advantage of context information collected by other applications, reducing its running time and energy consumption, memory and processing.

V. Framework Evaluation

To evaluate the CodeDroid framework we developed a context aware mobile application for a tourist scenario. In tourist cities that attract many visitors, it is very common to find a tour guide. The role of the guide is to help people to know the city, drive and visite the main sightseeing, recommend interesting places and local restaurants based on the profile, help visitors locate themselves in the city, among other functions.

The application *Places2Go*, a context-aware tour guide, was developed using the CodeDroid framework. The purpose of the application is to list points of interest located near the user. This application requires context information about the environment that is not provided by the CodeDroid framework. For getting information about locations in a given region, we used the Web service "*Search Venues*"

 ISBN: 978-1-61208-167-0

by *foursquare APIv2*[2]. The class diagram of the application Places2Go is depicted in Figure 1. The application classes are described below.

Place: entity responsible for storing information about places near the user. This entity has the attributes name, address, city, telephone number, category, distance, latitude and longitude.

Profile: represents the user's application profile. It is a subclass of the *ProfileBase* class of the CodeDroid framework. As expected, this subclass inherits all the attributes already defined for that profile and adds the attribute *Places*, a list of objects that represents the user's proximity locations.

ProfileController: acts as the control application and sends the commands to the responsible services. It is a subclass of *ProfileControllerBase* of the CodeDroid framework, inheriting the provided services already defined.

Places2Go: main application class that represents the initial screen. It is a subclass of *AcitivityBase* of the CodeDroid framework responsible for initializing the necessary services for the application in addition to those already started by the framework. It displays the current user's location and sets two buttons. The first one is the "*Update location*" to update the location information and the second one is "*Places*", which looks for a list of places near the user.

PlacesService: search service that looks for interest places in the region where the user is located. This service connects the Web service of the API *foursquare* and obtains the data related to the local region. It is a subclass of the Android's *Service* class.

ListPlaces: class responsible for displaying the list of places near the user. Each location is represented as a list item. For each one of them there is an associated icon with a visual representation of the place's category. Since this is a screen, it is a subclass of the Android's *Activity* class.

PlaceDetails: class that displays the details of a site. The fields are name, address, city, telephone number, category and distance. The latitude and longitude were omitted because they are not of interest to the user. Since this is a screen, it is a subclass of the Android's *Activity* class.

The location service is activated and deactivated through the *ProfileController* class, subclass of *ProfileControllerBase*. It was not necessary to define a new location service because the methods to access the service available at the framework CodeDroid were inherited from the *ProfileControllerBase* class. The application was implemented in such a way that whenever there is a change to the user's location to a distance greater than one kilometer, the location information is updated automatically.

The application displays the user's current location, and in case it is outdated, the user can select the "*Update location*". Upon detecting the location change, the application starts representing the new user's address. Then the user can select "*Places*" to know about suggestions for places nearby and, thus, the application returns the results for the query. Due to performance reasons, the user receives at most 50 suggestions. For each result, it is returned the name of the place, the distance between the user's location and the place, and an icon that represents the type of place. In some situations, the site was not properly categorized, so icons can be inconsistent, i.e., the application cannot take the responsibility for this kind of problem. On the screens of our example, illustrated in Figure 2, the user is located at Tiradentes Plaza in Ouro Preto[3], Minas Gerais. This is a landmark of the historic city. Around it, we have the most visited tourist sites. The application *Places2Go* returns important sites such as the Museum of Inconfidência, the Church of Nossa Senhora do Carmo, House of Tales, among others.

The application was successfully evaluated in two environments: the emulator using the Android platform and using the smartphone Samsung Galaxy S that runs Android version 2.1. The Android platform provides tools for evaluating the application. Among them, we use the feature that records the CPU usage. Figure 3 shows the CPU state in four moments that occurred during the application execution. The share of the application is identified by the name of the application package "br.ufmg.dcc.mestrado". We noticed that the CPU consumption in the worst case was equal to 11.12%. The CPU consumption was relatively low, showing that the application developed by the CodeDroid framework is a viable solution to be used on mobile devices.

VI. Conclusion and Future Work

This paper presented the concept of a profile as an entity to represent context, adding the most relevant information for decision-making applications. To demonstrate the use of the profile and help the development of context-aware mobile applications, we developed the CodeDroid framework. The application *Places2Go*, developed by the framework, has the capability of being extensible, modular and allows the reuse of components. The tool eases the development of applications, reducing the need for coding the generic functionalities provided by the CodeDroid framework. The application was evaluated in both emulated and real environments, with good performance results. As future work, we can mention the extension of the CodeDroid framework to support new features (e.g., sensing devives), the evaluation of the framework in different scenarios and its portability to other mobile platforms.

References

[1] Anind K. Dey, Gregory D. Abowd, Peter J. Brown, Nigel Davies, Mark Smith, and Pete Steggles. Towards a Better Understanding of Context and Context-Awareness. In *HUC*

[2]http://developer.foursquare.com/

[3]Ouro Preto is a world heritage city in Brazil. For further information please refer to http://whc.unesco.org/en/list/124.

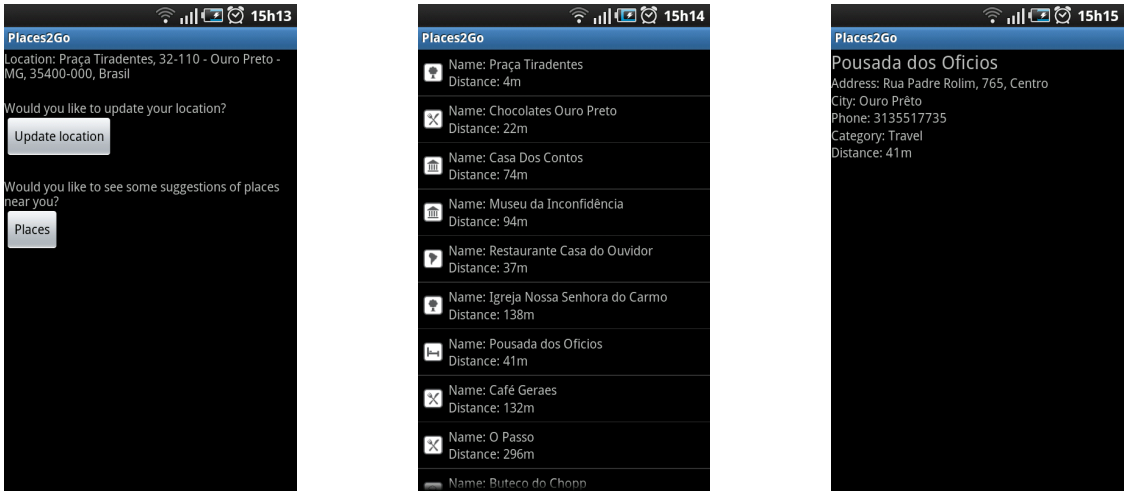


Figure 2. On the left, the application screen after clicking on *Places* (middle). On the right, the screen listing the places resulting from the action.

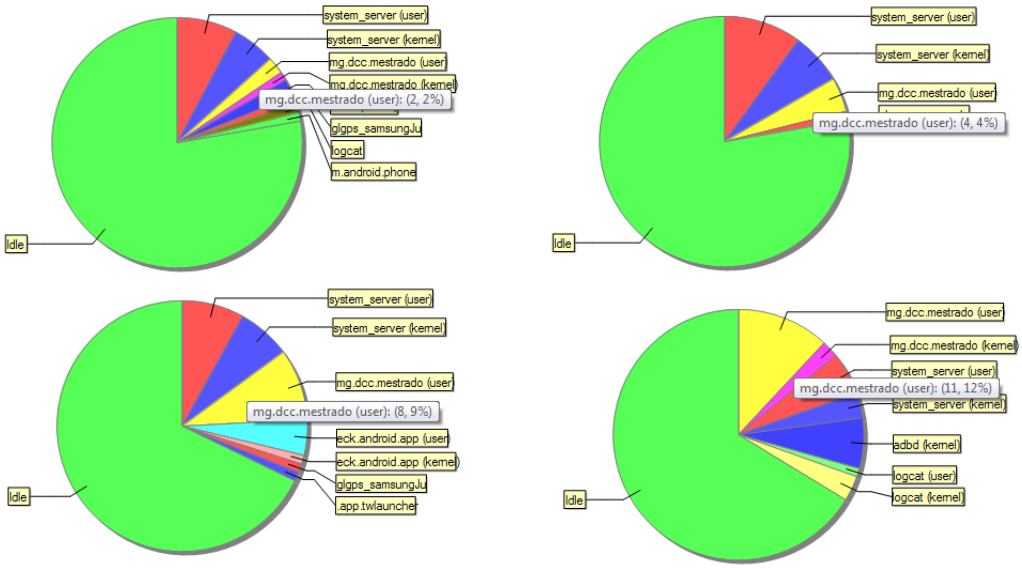


Figure 3. CPU usage of the smartphone running the application *Places2Go*.

'99: Proc. of the 1st Int'l Symp. on Handheld and Ubiquitous Computing, pages 304–307, 1999.

[2] Bart van Wissen, Nicholas Palmer, Roelof Kemp, Thilo Kielmann, and Henri Bal. ContextDroid: An Expression-Based Context Framework for Android. In *PhoneSense*, Zurich, Switzerland, 2010.

[3] Alf Inge Wang and Qadeer Khan Ahmad. CAMF – Context-Aware Machine Learning Framework for Android. In *IASTED International Conf. on Software Engineering and Applications – SEA*, Marina Del Rey, CA, USA, 2010.

 ISBN: 978-1-61208-167-0

A Personalization Approach Based on Models Integration for a User Interface for Supervision in a Power Plant

Victor Alvarez-Cortes
Gerencia de Supervisión de Procesos
Instituto de Investigaciones Eléctricas
Cuernavaca, México
vacortes@iie.org.mx

Benjamín E. Zayas Pérez
Gerencia de Supervisión de Procesos
Instituto de Investigaciones Eléctricas
Cuernavaca, México
zayas@iie.org.mx

Víctor H. Zárate Silva
Depto. de Tecnologías de Información y Mecatrónica
Instituto Tecnológico de Monterrey
Cuernavaca, México
vzarate@itesm.mx

Jorge A. Ramírez Uresti
Depto. de Tecnologías de Información Computación
Instituto Tecnológico de Monterrey
Estado de México, México
juresti@itesm.mx

***Abstract*— In the operation of a power plant, the operator has to deal with a number of variables displayed in several windows, presented in different formats such as tabular, one-line diagrams, process diagrams, etc. Under these conditions, the operator needs to locate data relevant to the current status of the plant and efficiently navigate a large number of screens within the user interface. This situation becomes critical for infrequent tasks within the plant such as the start-up, shut-down or for abnormal operational conditions. In order to deal with the above situations, we propose a personalization mechanism based on models for adapting contents in a user interfaces for the operation of power plants. The model takes advantage of the integration of specific sub-models from artificial intelligence, user modeling and human-computer interaction. A personalization approach is applied into an adaptive user interface prototype that is expected to improve the user-system interaction by reducing the time to complete the startup of a power plant. For this purpose, contextual information of the power plant, user interaction logs, user's preferences and experience are considered. At this first stage, our prototype was evaluated in a simulated scenario with non power plant operators to investigate improvements of task performance for carrying out the procedure of the startup of a power plant process and to identify usability issues. Results from this initial evaluation show consistent time reductions and correct predictions of futures displayed variables for the adaptive user interface version despite the fact that participants in this evaluation had a lack of the domain knowledge. This results needs to be validated with further studies involving power plant operators, in order to compare the impact the domain knowledge plays in the time reductions.**

Keywords: personalization; user interface; user modeling; adaptive interaction; user interface evaluation.

I. INTRODUCTION

The area of Intelligent User Interfaces (IUIs) is a specialized research field within the Human-Computer Interaction discipline (HCI), which pursues to overcome some of the limitations of traditional user interfaces. Traditional user interfaces for process supervision in power plants are comprised of a large number of displays such as tabular data, one-line diagrams, sequential diagram, historic trends, process diagrams, critical events logs, etc. The user needs to visualize information, analyze normal and failure situations, process changing data, understand the underlying process dynamics and finally take all this information into account to make a decision in a reduced time span. This amount of information represents a high cognitive load that the operator has to deal with and becomes critical for abnormal, failure or infrequent operations within the plant [1]. Established design process involves a user interfaces designer who establishes *a priori* the way each kind of data must be presented to the user and establishes a corresponding mapping between the processed information and the way to be displayed to the user. This mapping is achieved based on user interfaces guidelines, ergonomics studies, usability rules or accepted HCI standards; however, this mapping is static and the Graphic User Interface (GUI) is not designed for abnormal situations, where information flow is higher and operators have to face new situations [2].

II. PREVIOUS WORK

Current research efforts attempting to use adaptive user interfaces for critical domains include IUIs based on models, knowledge, examples and demonstration, plan recognition, task recognition, agents as well as multi-modals interactions. The challenges faced by the IUI community are diverse and disciplines such as HCI, Artificial Intelligence

 ISBN: 978-1-61208-167-0

(AI) and User Modeling (UM) have reported different research lines with specific approaches for IUIs.

Likewise, the generation, maintenance and use of user models as a core adaptation element has been another explored approach in recent years [3]. A search for novels ways to represent the knowledge a system has about its users, their skills, preferences and goals is another promising research line. Taking advantage of these user models, an IUI attempts to adapt its contents, layouts or navigation elements to suite the user experience or preferences [4]. Mixed-initiative interaction models have been developed to deal with the interruptions challenge and its timing while presenting information to the user [5]. Another approach followed in IUIs has been generating task models and use them as the central element of the adaptation strategy. The goal is to infer the next possible tasks and to anticipate actions the user is about to accomplish in order to find a way to assist her in the task at hand, or facilitate the use and learning of the user interface itself [6] [7].

Copious research efforts in the electrical domain have focused on the development of systems to help operators in fault situations [2], the design of intelligent systems to assist operators in normal power plant operation, and even more efforts in the development of complex systems to detect, predict and correct faults in real-time in power systems [8]. However, there is a clear lack of research in the area of adaptive user interfaces and personalization for critical domains, such as those found in power plants where unfrequented tasks and fault situations need to be handled by operators. This work is relevant since is one of the initial efforts where researchers adventure at proposing and applying an adaptive user interface beyond the laboratory prototype scope to handle real critical operations for process control.

III. PROBLEM STATEMENT

Context domain knowledge information is a valuable assessment that traditional direct manipulation interfaces do not take into account. Same situation applies to user experience and preferences; even though it is known empirically that these elements are important to operate a plant in a safety and optimal way [2]. It is our belief that in critical domains such as power generation plants, the use of the abundant existent contextual information and the user navigation history is valuable and its selective use in an adaptive user interface will improve the performance of human-computer interaction by means of personalization the variables displayed for infrequent or critical tasks.

Contextual information includes operators' preferences and domain knowledge, the tasks performed at the user interface, as well as the process information such as monitored variables, plant operation status and process stage, information already stored for process analysis purposes. By improving the performances we mean the adaptive interface will allow the operator to achieve specific operational tasks within a pre-defined time frame by filtering out irrelevant information and personalizing contents and navigation. Infrequent tasks accomplished correctly in sequence and time have a relevant impact, particularly to the startup and shut-down stages of a power plant operation where a deviation in time from a pre-established sequence increases the gas consumption and shortens the expected life span of the power plant.

IV. PERZONALIZATION APPROACH

For personalization purposes, we propose the use and fusion of several sub-models integrated to create an adaptive user interface that will improve the plant operation by personalizing information screens taking into account 4 key elements: (1) context information of the power plant, (2) the user-system interaction history, (3) user preference, and (4) experience.

The meta-model is composed of two set of models: (1) central models and (2) support models.

A. Central Models

These are the intelligent components that provide the adaptive behavior to the system by receiving and processing the information from the support models to provide adaptive information displays and navigation in a timely manner. To accomplish this goal, these components apply AI techniques to the collected information from support components.

1) RECOGNIZER

The RECOGNIZER model decides *when* to proceed to change the current interaction mode. Possible interaction modes includes: directed, initiated or controlled by the system, directed by the user or a combination of both. The RECOGNIZER uses a mixed-initiative approach and takes into account the interaction history and, by applying a utility-based algorithm, analyze possible alternatives and take the highest ranked option available to keep, or change if required, the interaction mode. The expected utility approach permits to establish the most convenient interaction mode to the current situation. Therefore, it determines if it has to keep with the current interaction mode, change to another interaction mode or, ask the operator for further data and interrupts the interaction. This decision has to be made based on the evaluations of advantages and disadvantages when interrupting the operator, and the known history for handling similar situations in the past. The information to decide includes the willingness of the user to interact with the system in the past, the willingness to personalize in the beginning. Once an adaptation is detected and ready to be presented to the user, the RECOGNIZER decides if the cost is high enough to make it available to the user or if a switch in the interaction mode has a higher cost as a distraction or interruption for the user (ranked by the rule generator). The RECOGNIZER was developed following a utility-based approach adapted from the research presented by Fleming in his work [9].

 ISBN: 978-1-61208-167-0

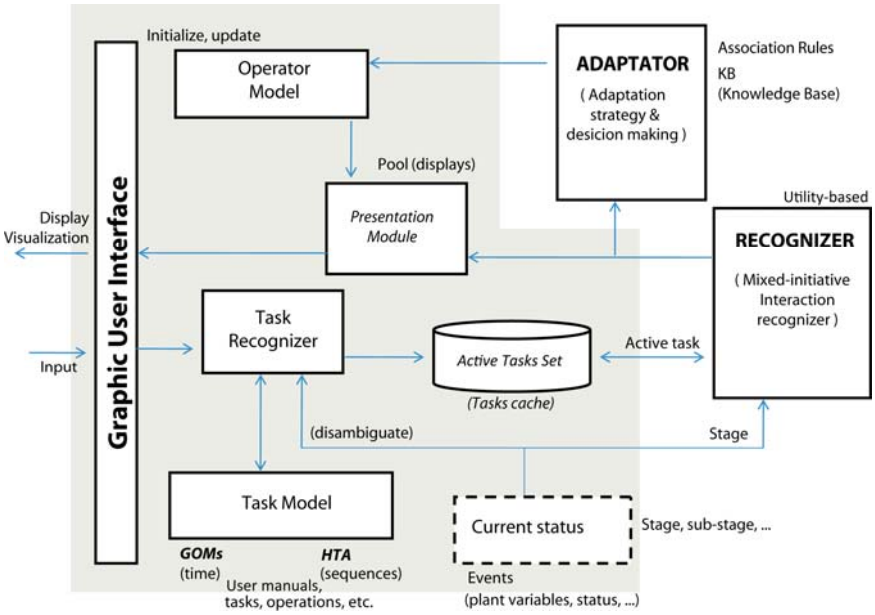


Figure 1. Central and support components for the proposed adaptive model

The costs of interruption computed in the RECOGNIZER are represented using a linear model, so the total cost is a weighted sum of any individual cost measures that have been identified for the application domain. For example, the identified factors for the startup of a power plant are: (1) user knowledge represented in the user model, (2) user willingness to interact determined by the initial personalization achieved by the user and the number of visited screens, (3) current context (current node for the task at hand), and (4) task performance expected due to interaction, provided by a domain expert. Each of these factors is normalized so that values range from 0 to 100. A cost of 0 indicates no cost for the interruption so the RECOGNIZER will allow to be presented the adapted content in the user interface. A cost of 100 is the maximum possible cost and in this situation the RECOGNIZER will not interfere to change the current presented content.

2) ADAPTATOR

This is a central component in charge of establishing the strategy for displaying contents and adapting the navigation to be presented to the operator. It determines what content will be presented as well as the way to be displayed. It also deals with data stored (history) for past interactions, contents and its presentation to the user along with the outcome of the prediction and the feedback received from the operator. We propose the use AI techniques, specifically applying Association Rules Mining (ARM) to generate a set of adaptation rules by borrowing some ideas from Bunt [7]. The ADAPTATOR was developed applying a modified AIS (*Agrawal, Ieminski, Suami*) algorithm using association rules mining techniques and extended the algorithm to support a meta-rule generator [10]. Rules are generated from process variables (i.e., pressure, temperature, gradient, dome level, etc.) correlated to previous interactions and the preferred variable presentations. A number of irrelevant and redundant rules are generated and in order to avoid its explosion we use domain expert knowledge to guide the meta-rule generation and to rank the generated association rules.

We describe a simple example to illustrate the application of our ADAPTATOR model regarding adapting contents to the user. The adaptive user interface has a set of default variables to be presented to the user for the current task. Data from support models such as user experience, display preferences, navigation traces, user willingness to interact and related plant process variables (i.e. temperature, dome pressure, temperature gradient, etc.) relevant to the current task are processed by the rule generator. Expert knowledge is used to eliminate redundant rules and also to guide the structure of the rules to be considered relevant. An example of a generated rule which includes experience level, visualization preference and process variable is presented in a simplified form as follows: IF Beginner AND Tabular AND Dome Temperature > 120 AND Dome Pressure Stable THEN Display = Set 3, where Set 3 is a predefined set of variables to be presented to the user. Note that the ADAPTATOR does not generate adaptive displays on-the-fly, but select one from a predefined pool of custom made sets. When this rule is generated, the specific set of variables (Set 3) is ready to be presented on the adapted area of the user interface.

B. Support Models

These components provide the information related to the tasks under supervision or control, the field variables of the power plant correlated to the startup stages, the operator's preferences and knowledge. These models are briefly described in the following sections.

 ISBN: 978-1-61208-167-0

1) Task Model

Its main function is to represent the tasks performed within the system considering its duration, as well as, its sequences. The operative knowledge about the tasks that are required to operate the plant through the user interface can be found in the plant manufacturer operation manual, operator's training manuals or can be extracted from observation of the user interface in action.

2) Operator Model

Based on an user *overlay* model, its main function is to manage and maintain the data associated to the operator characteristics, such as personal data, display and navigation information preferences, as well as her knowledge of the task to be performed. This model was developed with an overlay approach taking into account specified times for maximum, minimum and normal times for completing a task by a beginner, normal and expert operator during the startup of a power plant.

3) Task recognizer

It receives and integrates information regarding the task to be performed by the operator, and the possible available sequences. The task recognizer takes into account the operator's knowledge about the task available at this time and additionally integrates the operator's characteristics, which are retrieved from the operators sub-model. This module process the information to maintain a set of possible tasks similar to those proposed by Bunt [6]. These tasks are exchanged with the AST (Active Set Tasks) module as shown in Figure 1 above.

4) Presentation module

It is an isolated module separated from the central models that is constantly exchanging data with the central and support models. It is also perceived by the operator as the GUI (Graphical User Interface. A challenge for its design and implementation was the fact to be able to continuously logging navigation the traces during the interaction with the operator and save them to the knowledge database for its use by the rule generator.

V. ADAPTIVE USER INTERFACE PROTOTYPE

A prototype integrating the proposed models was developed. Due to the complexity of core components and the critical nature of supervision process of a power plant itself, the integration stage for this research project required unforeseen additional effort since different technologies and programming languages for different models were involved. Currently, we have finished the development of the ADAPTATOR and is also completed the testing of all the support models; however, the mixed-initiative RECOGNIZER is still under development, and its impact in the adaptive strategy is not researched in the evaluation presented in this paper.

Integration of all involved models was essential to proceed to the evaluation stage since its is required to collect the entire context information and navigation traces in order to feed the ADAPTATOR mining algorithms and correlate attribute-value pairs from the different models. The proposed meta-rule generator was also finished and is the key element to disambiguate and control the explosion in the number of association rules generated. The prototype for the Adaptive User Interface for Power Plant Startup (AUI-PPS) was the mean to implement the proposed mixed-initiative adaptive user model and is shown in Figure 2.

When operator uses the adaptive user interface prototype for the first time, it allows him to select his presentation preferences. Three GUI presentation modes are displayed for selection: minimalist, standard and graphical enhanced. These preferences are stored and updated in the operator model as nodes (tasks representation) of the startup process are completed.

Suggested presentations are displayed in a specific area (right side), which will be changing according to inferred data by the ADAPTATOR model, see Figure 2. User navigation traces are stored in order to integrate it with past interactions and user preferences and eventually decide whether to change the presentation or keep it until the expected utility function from the interaction RECOGNIZER reach specified trigger from the user model.

The selection of relevant variables to present, quantity and type of presentation for the task at hand are personalized constantly by the ADAPTATOR. These parameters are not changing all the time but when the RECOGNIZER algorithms evaluate to change the interaction mode as shown in Figure 3. The user can reject, at any given time, any of the suggested variables and its presentation, and his decisions are added into the navigation traces logger and feed backed to the mixed-initiative RECOGNIZER. By clicking on the red cross icon on the variable presentation space, the content is diminished and the user model is updated. For the RECOGNIZER and the ADAPTATOR this represents, an inaccurate prediction and is fed back to the central models for future content predictions in a similar situation.

VI. PROTOTYPE EVALUATION

For our evaluation, we selected a critical domain represented by the startup of a power plant, which is a complex process that takes hours (6-8 hours) and involves carefully operation of surrounding subsystems. For initial evaluations a proprietary simulator was used along with ASPPO (Aid System for a Power Plant Operation) that provided guidance to the operators during the startup of a power plant.

This initial evaluation was achieved with six users (non plant operators) with computer knowledge and related to the electrical domain, but without prior knowledge regarding the operation of a power plant.

A. Participants

The evaluation was carried out with a group of six subjects with different experience levels. For the evaluation, the participants had to visualize information and interact simultaneously with two LCD monitors: (1) to display the process diagrams of the plant simulator, where the user carried out the actions instructed by the adaptive and (2) a second monitor to display the adaptive interface with the

 ISBN: 978-1-61208-167-0

instructions to guide the participant in order to complete the current task.

Each participant had to evaluate both user interfaces (adaptive and non adaptive) with an assigned task of variable difficulty level.

B. *Evaluation Procedures*

A short introduction to the simulator and the startup process was given to the six participants. An introduction questionnaire was given to each participant in order to evaluate its domain knowledge and computational skills. At the end of the evaluation a final questionnaire was also applied to participants in order to capture the user's perception about the system and their general final thoughts.

All evaluation was achieved in an isolated room with two monitors: one for a simulator and another for the Adaptive User Interface Prototype for Power Plant Startup (AUI-PPS) in charge of displaying the recommended actions in order to achieve the task at hand required for completing the current procedure (node) of the startup stage.

Five of the most representative tasks for the pressurization stage and steam generator heating procedure were given to the participants. Taking into account the average time for completing each node for this stage, the 5 nodes should take 1 hour 43 minutes for an average user. Since this is a considerable time, it was harder than usual to find participants willing to spend so much time in an evaluation. The experiments followed an incremental approach (i.e., integrating additional models and evaluating its impact), so that the effectiveness and accuracy achieved by the adaptive user interface could be measured, isolated and attributed to specific models.

Figure 2. Adaptive user interface layout: adaptive content area on the right side and standard display area separated to minimize impact on user

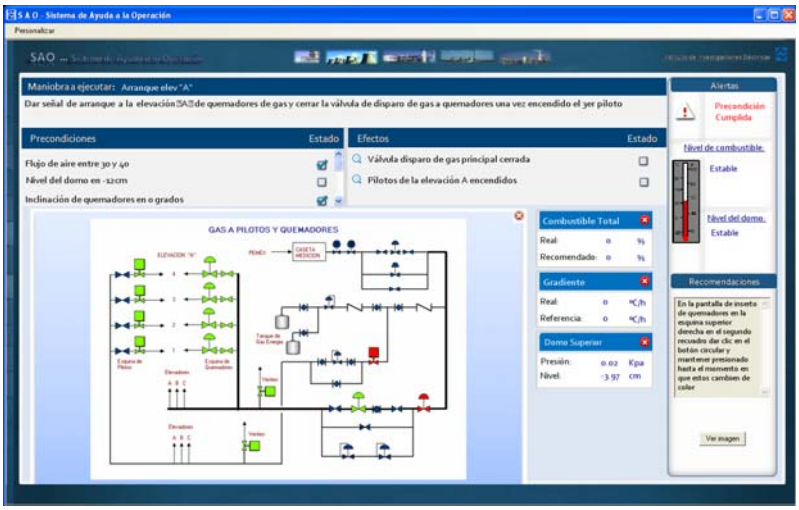


Figure 3. Adaptive user interface with recommendations presented in the lower right corner.

 ISBN: 978-1-61208-167-0

C. Experimental Evaluation

The experimental evaluation reported in this paper was achieved with six users using a prototype that integrated the operator model, the task model and the ADAPTATOR model. The Mixed-Initiative RECOGNIZER model was not included in this initial evaluation. The next step in this research work will include a greater number of participants, the inclusion of the RECOGNIZER and an evaluation with actual operators in a power plant in order to contrast and validate the initial results.

This experiment also included data provided by the task model in charge of detecting the currents task to accomplish (node) and its associated variables to be processed by the ADAPTATOR. The idea behind this incremental approach is to research the operator's behaviors and interactions improvements gained as each component is integrated into the adaptive model.

We use usability techniques to evaluate the adaptive prototype of the user interface [11]. This assessment provided us the base parameters in terms of efficiency, effectiveness as well as the operator's perception of the current non-adaptive user interface. A video recording and logging of interactions at the user interfaces level was achieved and analyzed.

VII. RESULTS

Results presented here are grouped in three key areas: efficiency (time, consumption), task completed and user's perception. It is important to note that this evaluation was carried out with participants with knowledge about computers but a lack of knowledge about the operation of a power plant. The rationale behind this approach is to study and measure the degree of impact the domain knowledge has (if any) in the completion of the tasks and its contribution to the overall time reduction in order to isolate it from the contribution of the adaptive user interface model itself. Our final goal will be, when finished the next evaluation stage with operators, to study the impact of the 2 different participants: (1) users with computational knowledge, associated with their ability and familiarity for handling a user interface and following the adaptive interface guidance to complete the task, regardless of their lack of domain knowledge and (2) participants (operators) with domain knowledge, associated with an in-depth knowledge of the start-up process and operation of a power plant.

Results of the evaluation of the operator model provided information regarding the user preferences and knowledge (or lack of) about the processes required for the startup of the power plant. This knowledge is used to adapt the presentation of the elements to achieve the operations recommended by the AUI-PPS. The user follows these recommendations and she needs to explore and find those components by navigating the different screens of the simulator user interface.

Figure 4 shows the times (in minutes) to complete the task assigned to the six participants, using the normal user interface, the adaptive user interface and the third bar shows the values predicted by the ADAPTATOR.

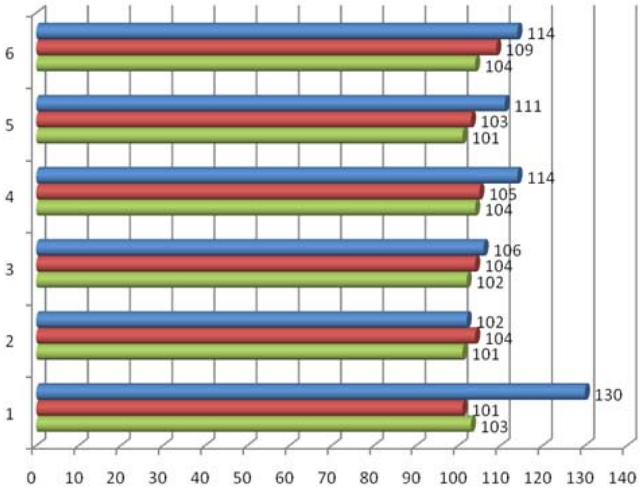


Figure 4. Time to complete the assigned tasks.

For participants 1, 3 and 5 the non-adaptive user interface was first evaluated. Participants 2, 4 and 6 first evaluated the adapted version of the user interface. In eache one of the evaluations consistently the user with the adapted version took less time to complete the tasks. The maximum difference was 29 minutes, however this was an isolated case, since the global average was 5.8 minutes.

Depending on his experience and knowledge so it is the time he takes to achieve the recommended task in the expected time. If the subject under evaluation follows the displayed recommendations and knows how to accomplish them on the simulator, he/she can finish the task on time, earlier, take longer or even abort the task by lack of knowledge or by misunderstanding the recommendations.

The time to complete the task to startup the power plant (pre-heat and pressurization of the steam generator) had a direct impact in the gas required to operate the plant as it is summarized in table 1.

TABLE I. GAS COMSUMPTION FOR BOTH USER INTERFACES

Subject	Non-adaptive User Interface	Adaptive User Interface	Predicted
1	9681.75 Kg.	7653.5 Kg.	7907.2 Kg.
2	7536.75 Kg.	7615.25 Kg.	7371 Kg.
3	7614.75 Kg.	8053 Kg.	7419.75 Kg.
4	8424.75 Kg.	7624.45 Kg.	6864 Kg.
5	8297.25 Kg.	7673.25 Kg.	7646 Kg.
6	7946.25 Kg.	7829.25 Kg.	7624.5 Kg.

 ISBN: 978-1-61208-167-0

The results from the post-test questionnaire show the user perception while interacting with both interfaces. Usability issues included 3 questions to capture perception for ease of use and learn for each user interface.

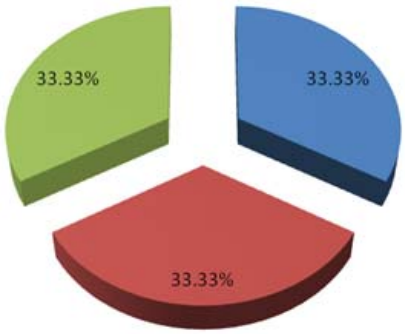


Figure 5. Balanced usability overall user perception for standard user interface: 33% easy, 33% neutral, 33% hard.

The adaptive user interface version is perceived more user friendly and easy to learn in general, however for some participants the user interfaces in both versions was perceived as "not easy" to learn. This perception was associated with the participant with less experience in both interfaces.

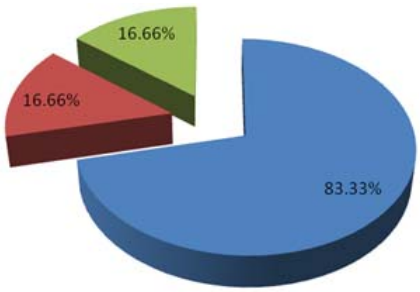


Figure 6. Usability overall user perception for adaptive user interface: 83% (easy to use and learn) and 16.66% for neutral and hard

VIII. CONCLUSIONS

Current results shows a consistent advantage of the adaptive user interface, by presenting personalized contents (variables and recommendations) over the traditional user interface for the startup of a power plant. A reduction in time of an average of 5.8 minutes for a reduced portion of the test of just 1 hour and 44 minutes that took the evaluated nodes is a significant reduction. If we keep the same linear proportion in the time reductions it is reasonably to forecast a reduction of 27 minutes for the whole process for the startup of a power plant, taking into account that this process normally takes 8-9 hours approximately. Likewise, this advantage shown in raw data is also present in the user perception about the usability of the adaptive interface.

These results seem to be promising for the adaptive user interface version versus the non adaptive version from the point of view of time reductions for this domain. However future work is required to include a greater number of participants in the next experiments and to carry out the evaluation with actual operators, so we expect that the difference in performance for both interfaces in the startup of a power plant might present different behavior.

REFERENCES

[1] Gianni Viano, Andrea Parodi, James Alty, Chris Khalil, Inaki Angulo, Daniele Biglino, Michel Crampes, Christophe Vaudry, Veronique Daurensan, and Philippe Lachaud. 2000. Adaptive user interface for process control based on multi-agent approach. In AVI '00: Proceedings of the working conference on Advanced visual interfaces, pages 201–204, New York, NY, USA. ACM Press. ISBN 1-58113-252-2.

[2] Zita A. Vale, A. Machado e Moura, M. Fernanda Fernandes, Albino Marques, Couto Rosado, and Carlos Ramos. 1996. Sparse: An intelligent alarm processor and operator assistant. IEEE Expert: Intelligent Systems and Their Applications,12(3):86–93, 1997. ISSN 0885-9000.

[3] Saleema Amershi and Cristina Conati., 2007. Unsupervised and supervised machine learning in user modeling for intelligent learning environments. Proceedings of the 12th international conference on Intelligent user interfaces, New York, NY, USA, ACM Press. ISBN 1-59593-481-2.

[4] Krzysztof Gajos and Daniel S. Weld. 2005. Preference elicitation for interface optimization. In UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology, pages 173–182, New York, NY, USA, ACM Press. ISBN 1-59593-271-2.

[5] Marcelo Armentano, Daniela Godoy, and Analia Amandi., 2006, Personal assistants: Direct manipulation vs. mixed initiatives interfaces.Int.J.Hum.-Comput.Stud.,64(1):27–35, 2006.ISSN1071-5819.

[6] Iain M. Begg, Joe Gnocato, and Wendy E. Moore.,1993. A prototype intelligent user interface for real-time supervisory control systems. In IUI '93: Proceedings of the 1st international conference on Intelligent user interfaces, pages 211–214, New York, NY, USA. ACM Press. ISBN 0-89791-556-9.

[7] Andrea Bunt, Cristina Conati, and Joanna McGrenere. 2007. Supporting interface customization using a mixed-initiative approach. In IUI '07: Proceedings of the 12th international conference on Intelligent user interfaces, pages 92–101, New York, NY, USA. ACM Press. ISBN 1-59593-481-2.

[8] Varanon Uraikul, Christine W. Chan, and Paitoon Tontiwachwuthikul. 2007. Artificial intelligence for monitoring and supervisory control of process systems. Eng. Appl. Artif. Intell., 20(2):115–131. ISSN 0952-1976.

[9] Michael Fleming and Robin Cohen. 2001. A utility-based theory of initiative in mixed-initiative systems. In Papers from the IJCAI-2001 Workshop on Delegation, Autonomy, and Control: Interacting with Autonomous Agents.

[10] Rakesh Agrawal, Tomasz Imielinski and Arun Swami. 1993. Mining association rules between sets of items in large databases. Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, May 25-28, ACM, New York, USA., pp: 207-216. ISBN:0-89791-592-5

[11] Janni Nielsen, Torkil Clemmensen, and Carsten Yssing. 2002. Getting access to what goes on in people's heads?: reflections on the think-aloud technique. In NordiCHI '02: Proceedings of the second Nordic conference on Human-Computer Interaction, pages 101–110, New York, NY, USA. ACM Press. ISBN 1-58113-616-1.

 ISBN: 978-1-61208-167-0