SMART 2012

The First International Conference on Smart Systems, Devices and Technologies

ISBN: 978-1-61208-225-7

May 27- June 1, 2012

Stuttgart, Germany

SMART 2012 Editors

Wolfgang Leister, Norsk Regnesentral, Norway

Petre Dini, Concordia University, Canada / China Space Agency Center, China

SMART 2012

Forward

The First International Conference on Smart Systems, Devices and Technologies (SMART 2012) held on May 27 - June 1, 2012 - Stuttgart, Germany, is an inaugural event covering tendencies towards future smart cities, specialized technologies and devices, environmental sensing, energy optimization, pollution control and socialcultural aspects.

Digital societies take rapid developments toward smart environments. More and more social services are digitally available to the citizens. The concept of 'smart cities' including all devices, services, technologies and applications associated with the concept sees a large adoption. Ubiquity and mobility added new dimensions to smart environments. Adoption of smartphones and digital finder maps, and increasing budgets for technical support of services to citizens settled a new behavioral paradigm of city inhabitants.

We take this opportunity to thank all the members of the SMART 2012 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 who dedicated much of their time and efforts to contribute to the SMART 2012. 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 are grateful to the members of the SMART 2012 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that SMART 2012 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in smart environment technologies.

We are convinced that the participants found the event useful and communications very open. The beautiful city of Stuttgart surely provided a pleasant environment during the conference and we hope you had a chance to visit the surroundings.

SMART 2012 Chairs
Petre Dini, Concordia University, Canada / China Space Agency Center, China

SMART 2012

Committee

SMART 2012 Technical Program Committee

Luis Alvarez Sabucedo, Universidade de Vigo, Spain
Ryosuke Ando, TTRI (Toyota Transportation Research Institute) , Japan
Petko Bakalov - Environmental Systems Research Institute, USA
Arati Baliga, AT&T Security Research Center, USA
Ranieri Baraglia, Italian National Research Council, Italy
Karolina Baras, University of Madeira - Funchal, Portugal
Stylianos Basagiannis, United Technologies Research Center Ireland -Cork, Ireland
Mehul Bhatt, University of Bremen, Germany
Rajendra V. Boppana, University of Texas - San Antonio, USA
Nidhal Bouaynaya, University of Arkansas at Little Rock, USA
Tone Bratteteig, University of. Oslo, Norway
Beatrice Cahour, National Centre for Scientific Research / Telecom ParisTech - Paris, France
Juan-Carlos Cano, Universitat Politècnica de València, Spain
António Casimiro, Universidade de Lisboa, Portugal
Hao Chen, State University of New York at Binghamton, USA
Liming (Luke) Chen, University of Ulster at Jordanstown, UK
Xin Chen, NAVTEQ Research, Nokia - Chicago, USA
Yao-Yi Chiang, University of Southern California, USA
Jaz Hee-jeong Choi, Institute for Creative Industries and Innovation, Australia
Diane Cook, Washington State University, USA
Antonio Coronato, National Research Council (CNR) - Napoli, Italy
Yang Cui, Huawei Technologies - Beijing, China
Rodolfo de Paz Alberola, United Technologies Research Center Ireland – Cork, Ireland
Vladimir Deart, Moscow Technical University of Communication and Informatics, Russia
Eugen Dedu, University of Franche-Comté, France
Jerker Delsing, Lulea University of Technology, Sweden
Catello Di Martino, University of Naples Federico II, Italy
Michahelles Florian, ETH-Zurich, Switzerland
Ahmed Frikha, University of Rennes I - IRISA, France
Kaori Fujinami, Tokyo University of Agriculture and Technology, Japan
Luigi Gallo, ICAR-CNR, Italy
Ivan Ganchev, University of Limerick, Ireland
Javier Manuela Gozalvez Sempere, The Universidad Miguel Hernández de Elche, Spain
Cornelia Graf, Center for Usability Research & Engineering, Austria
Bin Guo, Northwestern Polytechnical University, China
Jonna Häkkilä, Nokia Research Center - Oulu, Finland
Mohammad Hammoudeh, Manchester Metropolitan University, UK
Erik Hoel, Environmental Systems Research Institute, USA

Xiaodi Huang, Charles Sturt University, Australia
Xinyi Huang, Fujian Normal University, China
Kazuo Iwano, Smarter Cities - IBM, Japan
Ravi Jhawar, Università degli Studi di Milano – Crema, Italy
Marko Jaakola, VTT Technical Research Centre of Finland, Finland
Hoyoung Jeung, EPFL, Switzerland
Hanmin Jung, KISTI, Korea
Achilles Kameas, Hellenic Open University, Greece
Ibrahim Kamel, University of Sharjah, UAE
Alexander Klein, Technische Universität München, Germany
Christian Koch, Chalmers University of Technology - Gothenburg, Sweden
Lisimachos Kondi, University of Ioannina, Greece
Shin'ichi Konomi, University of Tokyo, Japan
Petri Koskelainen, Nokia - White Plains, USA
Kirill Krinkin, Saint-Petersburg Academic University / Russian Academy of Science, Russia
Ahmed Lbath, l'Université Joseph Fourier – Grenoble, France
John R. Lee, University of Edinburgh - Scotland, UK
Dan Lin, Missouri University of Science and Technology, USA
Juha Lindfors, University of Oulu, Finland
Linda Little, Northumbria University, UK
Giovanni Livraga, Università degli Studi di Milano - Crema, Italy
Veikko Louhevaara, University of Eastern Finland - Kuopio, Finland
Daniel Lucani, Instituto de Telecomunicações & University of Porto, Portugal
Artur Lugmayr, Tampere University of Technology, Finland
Xuanwen Luo, Sandvik Mining and Construction - Alachua, USA
Paul Lyons, Massey University - Palmerston North, New Zealand
Maode Ma, Nanyang Technological University, Singapore
Namunu Maddage, University of Melbourne, Australia
Tatiana Kozlova Madsen, Aalborg University, Denmark
Christian Makaya, Telcordia Technologies, USA
Yasir Malik, University of Sherbrooke, Canada
Yannis Manolopoulos, Aristotle University - Thessaloniki, Greece
Martina Mara, Ars Electronica Futurelab, Austria
Antonio Ruiz Martinez, University of Murcia, Spain
Sujith Samuel Mathew, University of Adelaide, Australia
Natarajan Meghanathan, Jackson State University - Jackson, USA
Harris Michail, Cyprus University of Technology, Cyprus
Carlos Miguel Tavares Calafate, Technical University of Valencia, Spain
Klaus Moessner, University of Surrey, UK
Robert Morelos-Zaragoza, San Jose State University, USA
Hongbo Ni, Northwestern Polytechnical University, China
Daniela Nicklas, University of Oldenburg, Germany
Petros Nicopolitidis, Aristotle University of Thessaloniki, Greece
Chris Nugent, University of Ulster, UK
Hadi Otrok, Khalifa University of Science, Technology & Research (KUSTAR) - Abu Dhabi, UAE
Luigi Patrono, University of Salento, Italy

Jun Peng, University of Texas - Pan American- Edinburg, USA
Fernando Pereñíguez García, University of Murcia, Spain
Fabio Pianesi, Fondazione Bruno Kessler - Trento, Italy
Christian Prehofer, Kompetenzfeldleiter Adaptive Kommunikationssysteme / Fraunhofer-Einrichtung für Systeme der Kommunikationstechnik ESK - München, Germnay
Juwel Rana, Luleå University of Technology, Sweden
Sherif Rashad, Morehead State University, USA
Arto Reiman, University of Oulu, Finland
Chiara Renso, CNR - Pisa, Italy
Joel Rogriques, University of Beira Interior/Institute of Telecommunications, Portugal
Kouchi Sakurai, Kyushu University-Fukuoka, Japan
Susana Sargento, University of Aveiro/Institute of Telecommunications, Portugal
Panagiotis Sarigiannidis, University of Western Macedonia - Kozani, Greece
Trenton Schulz, Norwegian Computer Center (Norsk Regnesentral), Norway
Rich Simpson, University of Pittsburgh, USA
Jeffrey Soar, University of Southern Queensland, Australia
Karthik Srinivasan, Infosys Labs - Bangalore, India
Young-Joo Suh, Pohang University of Science and Technology (POSTECH), Korea
Andrei Sukhov, Internet TV Ltd., Russia
Kåre Synnes, Luleå University of Technology, Sweden
Toshiyo Tamura, Chiba University, Japan
Markus Taumberger, VTT - Technical Research Centre of Finland, Finland
Dian Tjondronegoro, Queensland University of Technology, Australia
Nico Van de Weghe, Ghent University, Belgium
Arno Wacker, University Kassel, Germany
Alexander Wiethoff, University of Munich, Germany
Ouri Wolfson, University of Illinois - Chicago, USA
Qin Xin, University of Faroe Islands, Demark
Ying-Qing Xu, Tsinghua University - Beijing, China
Zheng Yan, Aalto University, Finland
Guanghua Yang, The University of Hong Kong, Hng Kong
Mao Ye, The Pennsylvania State University, USA
Jin Soung Yoo, Indiana University - Purdue University - Fort Wayne, USA
Jing Yuan, University of Science and Technology of China, China
Wuyi Yue, Konan University, Japan
Sherali Zeadally, University of the District of Columbia, USA
Chengyang Zhang, Teradata Inc., USA

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

Design and Development of Interactive Mirror for Aware Home

Chidambaram Sethukkarasi, Vijayadharan SuseelaKumari HariKrishnan, Raja Pitchiah
National Ubiquitous Computing Research Centre
Centre for Development of Advanced Computing
Chennai, India
{ctsethu, harikrishnans, rpitchiah}@cdac.in

Abstract— **This paper describes the design, development and deployment of an "Interactive Mirror", an artifact augmented with intelligence to demonstrate personalized services for enhanced comfort. The mirror aims at recognizing the user based on image processing techniques and provides personalized services like emotion recognition, health progress representation, event reminder and mirror usage time. The prototype of the interactive mirror was deployed in laboratory environment and the user's feedbacks were obtained. The face recognition and emotion recognition algorithms were tested and the results are discussed.**

Keywords - Ubiquitous Computing; Interactive mirror; Face Recognition; Emotion Recognition; Smart Artifact.

I. INTRODUCTION

Ubiquitous computing, with wirelessly connected embedded devices that are being used in day–to–day activities, is changing and improving our quality of life. Based on ubiquitous computing and communication technologies, many devices/products are emerging to make users comfortable. We daily look at the mirror and interact with it psychologically to find out how we look and how our attire is. Generally everyone spends some time to select a suitable outfit for the day during mirror usage. The interactive mirror is a development effort to augment the mirror with embedded intelligence for offering enhanced features such as face recognition, facial expression analysis, health progress reporting and selection of appropriate dress for the day. The 'Interactive mirror' is formulated as follows

Interactive Mirror = Mirror + Sensors + Database + Intelligence + Display

The Interactive Mirror comprises of a dielectric coated mirror mounted over a LCD Display, a camera for capturing the user's image, load sensors for measuring user's weight, Radio-frequency identification (RFID) reader and RFID tags for identifying the garment worn by the user. The mirror gets the inputs from all these sensors, processes it and provides output by displaying text and graphics as well as audio output.

Our contributions are as follows: 1) Conceptualization, design and development of interactive mirror that provides personalized services based on biometric identification of the user. 2) Facial feature extraction for emotion recognition 3) Training face recognition and emotion recognition algorithm with the database images created by us 4) Test results of face recognition and emotion recognition algorithm in the deployed environment.

The paper is organized as follows. Section 2 comprises of our comments about the related work, Section 3 details about the engineering of interactive mirror, Section 4 gives an overview of our proposed system and different modules involved in it, Section 5 presents the results obtained, and finally, Section 6 concludes the work.

II. RELATED WORK

Few investigations have been done in this area. By keeping the mirror usage in mind, certain efforts have been done to add the technologies in the mirror to do multiple tasks parallel at a time. How it will be if the mirror is talking to us. This concept has been realized in [2] using speech processing techniques. The mirror interacts with the user through verbal commands, functions like a good friend, listens to user's question and responds them, provides relaxation and consolation. It is essential to do the tasks parallel in our daily activities to save our time. Philips laboratory has come out with such a mirror [3] to assist the users in saving their time. One can watch news reports, TV channels, weather reports, etc., while brushing in bathroom. The bathroom lighting comprises of 50 light sources of different kind. Various light sources have been used which generates light of different color and temperature.

A mirror [4] makes use of behavioral data in order to provide its user with continuous visual feedback on their behavior in a natural manner. It tells you how you will look after 5 years.

An augmented setup [5] has been demonstrated targeting card game application. The system captures the image of the surrounding, detects multiple objects on it and provides augmented display on a LCD screen mounted on a table.

A digital mirror that detects tracks and model the human face and expression on a computer screen is demonstrated in [6]. The system uses robust multiple face detector and tracker based on active infrared (IR) illumination, and developed a physics based face model to generate realistic graphics output, and tested the integration of both modules using an eye-contact application, that randomly changes facial expressions.

Our system assists in leading a healthier life by measuring and providing health progress of basic health parameters, in addition to face recognition and facial expression analysis. The system also attempts to keep track

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

of the garments used over a period of time and suggest an appropriate dress for a particular day. Even though the system is designed for user without any disabilities, visually impaired people might find it more useful.

III. Engineering of Interactive Mirror

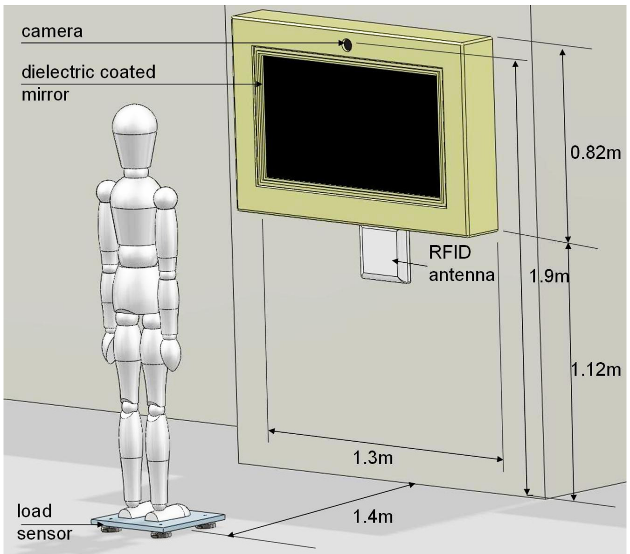


Figure 1. Engineering of Interactive Mirror

The dielectric coated mirror of thickness ¼" is mounted over the LCD display, so that the entire panel acts both as a mirror and display. When the display is ON, the dielectric glass will be transparent and when display is OFF, it will act as a normal mirror. A Creative web camera having VGA resolution and frame rate of 30fps is mounted on the top of the panel for capturing the images. The camera will be on sleep mode and triggered by the application for capturing mode when the user starts using the mirror. A weight measuring platform is formed with four load sensors from Loadstar. Each load sensor can carry weight up to 50lb, resulting in the maximum weight of the platform is 200lb. The weighing platform and the camera are mounted at suitable distance so that the entire face image should be covered by the camera Field of View (FOV). The distance between the weighing platform and the camera is kept constant to avoid scaling changes in the face image. The entire application has to be ported on atom processor which is placed inside the top frame of the panel. The omni directional, circular polarized Poynting UHF RFID antenna having 7dbi gain is positioned under the panel to detect the RFID tag attached in the garments. The antenna is connected to the SIRIT Infinity 510 RFID reader that is placed inside the wooden frame. The entire setup is shown in Figure 1.

IV. System Overview

The system encompasses of several modules to implement the services shown in Figure 2. The modules are described as follows.

A. *Face Recognition*

The steps involved in face recognition module are shown in Figure 3. The image captured by the camera needs to be preprocessed in order to enhance the contrast. Generally, an image is a collection of pixels. The pixel value refers to the intensity. To improve the image contrast, the intensity value is analyzed from top left pixel to bottom right pixel. The intensity value which occurred frequently is distributed over the entire image using histogram equalization technique.

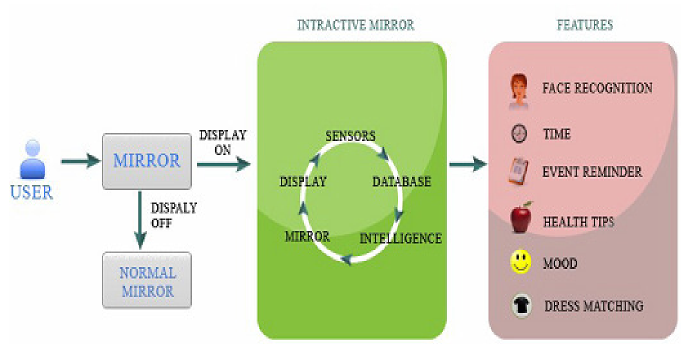


Figure 2. Services of Interactive Mirror.

The first step of face recognition is face detection. The process of finding faces in an image is called as face detection. Processing an image, pixel by pixel is a time consuming process. The methods available for face detection and face recognition are surveyed and compared in [7]. The system is using Viola and Jones algorithm [1] for face detection. It processes the image based on haar-like features and not pixels. The image is scanned from top left to bottom right using a window. The scan looks for the presence of haar-like features. The scan window is passed through a chain of filters called as classifiers. The classifiers are all trained with both face and non face images. The window which passes through all the classifiers are classified as "face" and if a window is rejected by any of the classifiers in the chain is classified as "Non-Face".

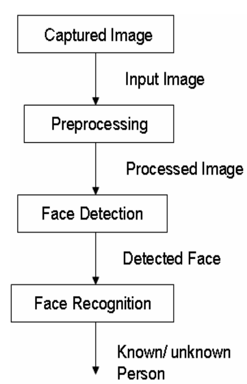


Figure 3. Steps involved in Face Recogntiion module.

The second step is to recognize whose face is this? The detected face image is to be cropped and resized for further processing. Many algorithms exist for face recognition

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

technique [13]. The face recognition technique using geometric approach and template matching have been compared in [12]. “EigenFace” Recognition method is used here [8][9][10] taking advantage of the actuality that the faces images will be frontal and upright while the user uses the mirror. This method works on training basis. The algorithm includes projecting the images over a lower dimensional space and finding the distances between them.

The face images of the users are captured and stored in a face database after processing shown in Figure 4. The algorithm needs to be trained with these standard face database images.

Figure 4. Face DataBase Images.

As said earlier, processing an image pixel by pixel is a time consuming process. For example, an image of size 100 * 100 is said to be a 10000 dimensional image. Hence, in order to reduce the dimension, principal component analysis (PCA) technique has been used. The Eigenvalues and Eigenvectors for all the images are calculated and projected over a lower dimension space.

Then the distance between the input image and all the training images in the projected space needs to be calculated using “Euclidean Distance” method. The Euclidean distance between two point’s p1(x1, y1) and p2(x2, y2) is given by (1),

$$d = \sqrt{((x2 - x1)^2 - (y2 - y1)^2} \quad (1)$$

Based on the distance measured, the input image is recognized. If the input face is a known face, the user is authorized, otherwise the user is unauthorized. OpenCV, an image processing library developed by Intel Corporation has been used for image processing algorithms.

B. *Emotion Recognition*

Generally, a person’s emotional state is very important, since it has effect for whole day. Human Computer interaction by understanding the human’s emotion is a recent research topic. Interactive mirror detects the user’s emotions based on their facial expressions while the user is interacting with the mirror. There exist various methods to recognize the user’s emotion. However, the system infers the emotions from the facial expressions as it get the input as face image. Facial expressions can be recognized from the facial features as per the steps shown in Figure 5. Facial features are the spatial position of the features and its displacements.

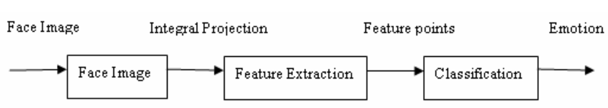


Figure 5. Steps inviolved in Emotion Recognition Module.

The following assumptions have been made 1) Constant lighting and illumination 2) Front Face image 3) No scaling variations.

The mouth features best describes the emotion of a person. Hence, mouth features like mouth width and height have been taken into consideration. Eyes also play a vital role in finding the emotion of the person. Thus eye features like eye width and eye height have also been extracted. The facial expression database images for happy, surprise and sad are shown in Figures 6, 7 and 8, respectively.

Figure 6. Happy Expression Database Images.

Figure 7. Surprise Expression Database Images.

Figure 8. Sad Expression Database Images.

Integral Projection method is used to extract facial features [11][14]. This method is affected by the external conditions like illumination changes, and skin color.

1) Mouth Feature Extraction:

In order to extract the mouth features, the search region has to limited by segmenting the mouth part alone from the face image. There exist three approaches [11] to extract the face features (i) based on luminance, chrominance, facial geometry and symmetry [12] (ii) template matching [15] (iii) PCA. Facial geometry approach is being used in our application. The steps involved in extracting the mouth features are shown in Figure 9. As per the face geometry, the mouth region is segmented from the face image. The segmented image is preprocessed to improve its contrast and converted to binary image using image thersholding. The threshold value is selected dynamically by the algorithm as per the image brightness. The unwanted blobs as shown in Figure 10a) occurred because of illumination variation along the face region is eliminated using blob removal technique. The blob removed image is shown in Figure 10b). A minimum blob size and the connectivity of the pixels are analyzed to say whether the blob belong to mouth region or not. By applying the integral projection method on binary

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

image, the exact position of the mouth corners and top point is located.

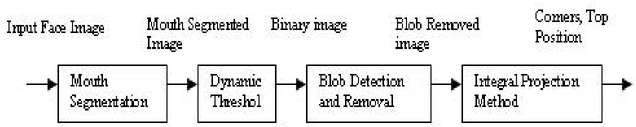


Figure 9. Steps involved in Mouth Feature Extraction

Let I(x, y) be a gray value of an image.
The Horizontal integral projection [11] of the image m x n is defined as

$$H(y) = \sum_{y=0}^{n} I(x, y) \tag{2}$$

The Vertical integral projection of the image is defined as

$$V(x) = \sum_{x=0}^{m} I(x, y) \tag{3}$$

a) b)

Figure 10. a) Binary Image b) Blob removed Image

where m = no of rows and n = no of columns.

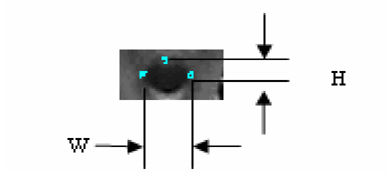


Figure 11. Mouth Feature Points.

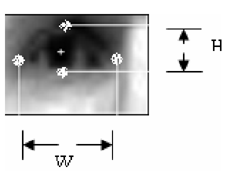


Figure 12. Eye Feature Points

The minimum value observed in the horizontal projection curve is the centre line of the upper and lower lip. Scan along the line both left and right to locate the corners. In vertical projection, the minimum value is the centre line of the mouth and scanning from top to bottom locates the top and bottom point of the mouth respectively. The mouth width is obtained from the two corner points and mouth height is obtained from the top point and the centre point of the mouth centre line shown in Figure 11.

2) Eye Feature Extraction:

As like mouth image segmentation, the same steps are followed for eye feature extraction also. The left and right eye region is separated from the face image as per the face geometry. The integral projection method is applied to locate the feature points as shown in Figure 12. The eye width and height is calculated from the featured points extracted. The minimum value in both the projection curve denotes the iris centre point.

3) k-NN Classifiaction Algorithm:

There exist various classification algorithms [18] suitable for facial expression recognition; the system make use of k-NN classifier. Similar works using k-NN classifier are available at [17]. It is one of the simplest machine learning algorithms. An object is classified by the majority vote of its neighbors. K is a constant denotes the no of neighbors. The value k should not be too small and not too large. The different k values are used for testing the accuracy of our algorithm and the results were discussed in section IV. The classification algorithm is trained with the extracted mouth feature values for classifying emotions like happy, surprise, sad and normal.

C. Health Progress Representation

The basic health parameters like weight, height, BMI and BMR are measured and saved in the health database. The values measured are saved in the health database along with the date and time of measurement. The health database is secured so that only the authorized person can have access after proper authentication.

The weighting platform mounted in front of the mirror starts giving the weight output when the user stands over it. The height information is obtained from the database currently; later the system will be integrated with the ultrasonic sensor to obtain height information of the user. Based on the weight and height information, the parameters like Body Mass Index (BMI) and Basal Metabolic Rate (BMR) have been calculated using the following formulas shown in equation 4 and 5/6 respectively.

$$BMI = (Weight) / (Height)^2 \tag{4}$$

Weight in Kilograms and Height in meter.
For Women,

$$BMR = 655 + (4.35 x Weight) + (4.7 x Height) - (4.7 x Age) Kcal / day \tag{5}$$

For Men,

$$BMR = 66 + (6.23 x Weight) + (12.7 x Height) - (6.8 x Age) Kcal / day \tag{6}$$

Weight in pounds, height in inches and age in years.

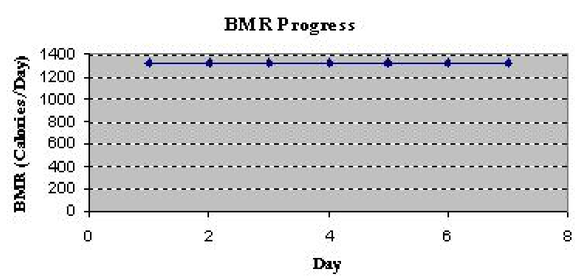


Figure 13. BMR Progress Representation.

BMI determines a person's weight according to his height and BMR is the measure of number of calories burned by the body when he/she is at rest. This helps in determining how much calories he need to intake inorder to maintain an energy balance and balanced diet condition.

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

Both of these parameters are important to measure and monitor to ensure that the person is healthy. The measured BMI value is analyzed for the conditions such as normal, overweight and underweight.

Drastic weight change over a short time period is the main symptom of identifying certain major diseases in human body. Hence the parameters measured have to be analyzed to monitor the change over a certain time period. Therefore the measured values are retrieved from the database and represented graphically to the user inorder to analyze any drastic change has occurred or not. For example, the BMR progress is shown in Figure 13. Similarly, the progress in weight and BMI are also analyzed and displayed to the user. The weight measurement accuracy depends on the accuracy of load sensor.

D. Event Reminder

In this busy world, it's normal to forget certain things like bill payments, birthdays, and important dates etc., hence, there is a need to have reminders to remind all those things. Interactive mirror will do for us. It reminds the important dates in the calendars, birthday, bill payments, tour date, etc.

The user has to enter the details which they want to be reminded using a GUI. When the user is using the mirror, the system will check for reminders against him/her at that particular time. If so, it will remind and also send a message to a hand held device of the user

E. Garment Description and Suggesstion

This feature is proposed to be implemented. The system makes use of RFID technology to identify the garment worn. The RFID tags are attached to the garments. The tag is flexible, water and heat resistant.

When a tag comes into the antenna coverage area, the reader detects the tag. Once the tag id detected by the reader, the details like when it was purchased, where purchased, price, color, material, how many times worn etc are retrieved from the database and displayed to the user. The suitable dress for the particular day is suggested by the mirror according to the presence of events like marriage function, birthday party, etc.

In textile application, the mirror can suggest a suitable dress color based on the skin tone of the user. The first step is to segment the skin color pixels from the captured face image. Next, the skin color pixels are analyzed with a certain predefined values to find the skin tone of the user.

RGB color space is used for skin color segmentation. The face image captured by the camera is scanned pixel by pixel from top left to bottom right. The pixels values are analyzed to classify whether the pixel belongs to skin color or not.

The following condition is used for classification.

A RGB pixel is classified as a skin color pixel, if the following condition is satisfied, the pixel is said to be skin color pixel.

$$R > 95 \,\&\&\, G > 40 \,\&\&\, B > 20 \,\&\&\, (\max\{R,G,B\} - \min\{R,G,B\}) > 15 \,\&\&\, |R-G| > 15 \,\&\&\, R > G \,\&\&\, R > B$$

Else the pixel doesn't belong to skin color.

Once the skin color regions are segmented, the values are analyzed and based on the values and the dress color availability with the particular user, a suitable dress is suggested for the particular day.

F. Mirror Usage Time

Interactive mirror sense the amount of time an individual uses the system and signals them if they are using for a long time. Based on the face recognition, the user's profile is unlocked and their school/office is checked. It tells how much time is remaining for them to get ready. If they have enough time to get ready then make up tips will be provided.

The weight measuring platform is used to find out the mirror usage time. The weight over the platform is monitored continuously. When anyone starts using the mirror, the weight will go up. That gives the start time. When the user leaves the load sensor platform, the weight will gradually decrease and then goes to zero and that time gives the end time. The time difference gives the mirror usage time.

Figure 14. Deployment of Interactive Mirror in Ubicomp Laboratory.

The demo illustration of all these features in the interactive mirror is represented in Figure 14.

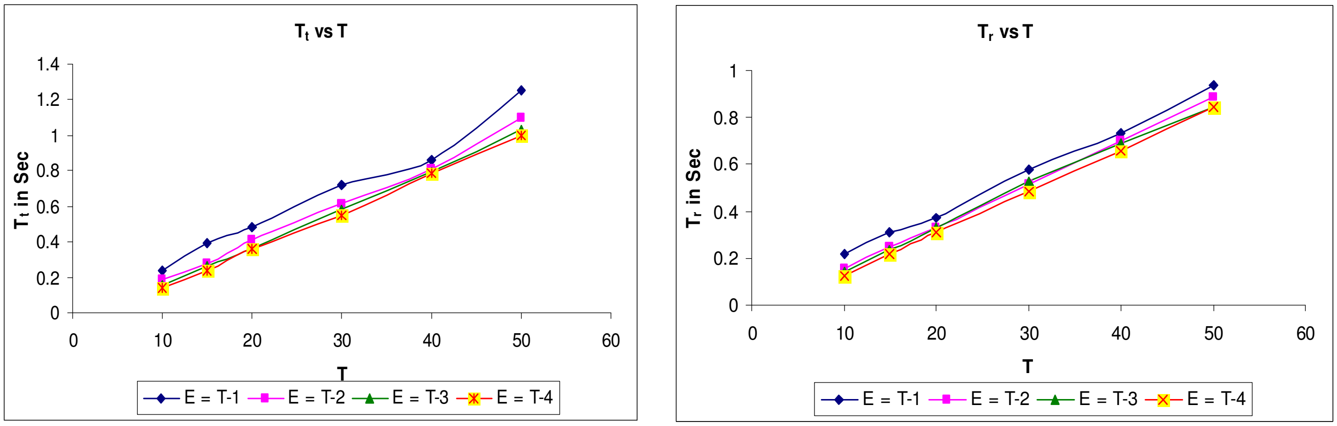
V. RESULTS

A. Face Recognition Algorithm Accuracy in Deployed Environment

The face detection algorithm had been tested with CMU face database. The false detection percentage and missing face rate were found to be 7.56 and 9.18, respectively. For example, the background partition glass image (non-facial image) shown in Figure 16 was detected as face image by the face detection algorithm.

Assume that the face images were frontal face images with no scaling variations taken under constant illumination. The users were asked to look at the camera and no other restrictions for them. There was no special experimental setup for capturing face image like dedicated light settings, user positioning with respect to camera, user's viewing angle on a particular object, etc. This created a real life testing environment. Therefore our database images are not strictly front face images. The face images may be tilted, pose varied, with/without expressions, etc., but not with much

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7



a) b)

Figure 15. Training time and Recognition time versus number of training images for different Eigenvalues a)Training time b)Recognition time.

variation. The training set consists of 42 images of 9 different subjects. The test set consists of known faces, unknown faces and non face images (false positives). Mahalanobis distance [19] method is better to use than the euclidean distance method. Inorder to recognize the unknown images and non face images, a proper threshold values have to be set for the distance calculation. These threshold values have to be calculated by experimentation for the chosen training set. Two different cases were considered for recognition. The one was to recognize the test image based on the first match alone and other based on the majority of first 3 matches. The accuracy of the algorithm for these two cases is listed in Table I. The role of number of eigenvalues (E) used in determining the algorithm accuracy was tested. It is noted that the accuracy decreases with number of Eigenvalues.

Among the two cases, the accuracy was found to be better in the case of considering the first match alone for recognition.

TABLE I. ALGORITHM ACCURACY FOR DIFFERENT NUMBER OF EIGENVALUES

Case	E = 41	E = 40	E = 39	E = 38	E = 37
First Match alone	91	77	72	71	68
Majority of first three matches	53	39	27	36	23

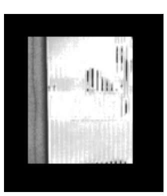


Figure 16. False Positive Image

The false positive result shown in Figure 16 had been eliminated out in face recognition algorithm as “Non face” image by setting up the proper threshold value.

B. *Face Recogntiion Algorithm Processing Time*

The training time (T_t) and recognition time (T_r) of the algorithm were obtained for different number of training images (T) and E. The results are represented in Figure 15a) and b). It was observed that both training time and recognition time increase with respect to number of training images. More number of training images leads to larger face space which in turn increases the training and recognition time. However the training time doesn’t affect the application much since the training has to be done only once at the initial stage. Later recognition process alone will run repeatedly.

The role of number of eigenvalues used in determining the algorithm processing time was also examined and presented. It was observed that the fewer number of Eigenface reduces the training and recognition time.

For achieving better accuracy, we were using maximum number of Eigenvalues.

C. *Emotion Recognition Algorithm Accuracy*

The Emotion recognition algorithm had been tested with our database images. The algorithm was trained with the mouth features like mouth width, mouth height, and left mouth corner to top point distance and right mouth corner to top point distance. The algorithm was tested with Yale facial expression database for different number of nearest neighbors (K). The results are presented in Table II. The algorithm was capable enough to recognize the happy expression when compared to other expressions. The extracted eye features has to be also included in the classification algorithm to improve the algorithm accuracy. The algorithm has to be tested in the deployed environment.

TABLE II. EMOTION RECOGNITION ALGORITHM ACCURACY

K	Happy	Surprise	Normal	Sad
3	93.33	53.33	66.66	33.33
6	100	60	46.6	26.66

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

TABLE III. USER'S FEEDBACK ON INTERACTIVE MIRROR SERVICES

	Face Recognition		**Emotion Recognition**		**Health Progress Representation**		**Event Reminder**		**Mirror Usage Time**	
	Rating	**Comments**	**Rating**	**Comments**	**Rating**	**Comments**	**Rating**	**Comments**	**Rating**	**Comments**
User 1	4	No comments	3	Should be Improved	4	No comments	5	No comments	4	
User 2	3	Sometimes the algorithm has reported as unknwon face, eventhough the information is already in database.	4	More emotions can be added to the database	4	No comments	4	How to input events to the system? Can we have some interaction where the suer can give inputs as well?	4	No comments
User 3	3	Sometimes good	3	Sometimes good	4	Good	1	Am not using yet	5	Good
User 4	3	Almost 60% it is Giving correct result	3	Same comments here also	1	Not at all watching, you may make it colorful and attractive	1	Didn't use till now, so don't know its capability	2	Not very sure how much it is getting reduced
User 5	3	Good only the environment is similar to training. Sometimes recognizing wrongly	3	Sometimes showing wrongly	4	Good	1	I haven't used it yet	5	Good
User 6	4	Mostly recognizing correctly. Problem arises only when lighting varies	2	Difficult part	5	Good		NA	5	Good
User 7	3	Still need complex algorithm of two or more algorithms running parallel for recognition	4	Eye and nose can also be considered in addition to lips for mood recognition	2	Instead of plain graph, info-graphics can be used		NA	3	No comments

D. *Weight Measurement Accuracy*

The weight measurement accuracy depends on the load sensor accuracy. The load sensor accuracy was tested with standard weights. The percentage error for various loads on a 50lb load sensor is shown in Figure 17.

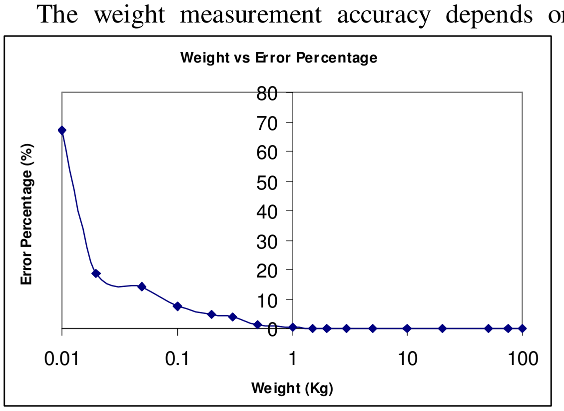


Figure 17. Weight versus error percentage.

The test results conclude the following 1) The error percentage is very high at 10grams, gets decreased up to 100g and finally very low and got stabilized after 1 Kg. 2) The load sensors should be loaded with maximum capacity to reduce the error percentage.

The rest of the features like mirror usage time and event reminder were found to be useful.

E. *User Evaluation of the System*

The system is deployed in our UBICOMP laboratory. The feedback has been obtained from the users (age group 24 to 32) for the services such as face recognition, emotion recognition, health progress representation, event reminder

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

and mirror usage time and listed in Table III. The rating has been done from 5 to 1 based on comfort/convenience level achieved in each service where 5 is the highest rate represents more comfortable and 1 represents not at all comfortable. The suggestions for improvements are being analyzed and suitable modifications will be carried out in the future development work.

VI. CONCLUSION AND FUTURE WORK

The Interactive Mirror system has been designed developed and deployed in our Ubiquitous computing laboratory. The features are all personalized using face recognition technique. The health parameter measurement and analysis help the user in leading a healthier life. The computation time and accuracy of the face recognition algorithm have been analyzed and discussed. The accuracy of emotion recognition algorithm is tested and results are presented. The system is targeted towards smart home application. It can be also used in Beauty Parlors, Textile Shops, and Hotels with some modifications in the offered features. The new user is trained automatically by the system when they use first time.

The system can be made much more useful to the users by adding more functionality like integrating light settings, speech processing, etc. The accuracy of the emotion recognition algorithm can be improved by including the extracted eye feature. The system can be deployed in real life home settings and feedback from wider set of users could be obtained for further enhancement of the system.

ACKNOWLEDGMENT

The work is developed under National Ubiquitous Computing Research Centre, funded by Department of Information Technology, Government of India. We would like to thank Department of Electronics and Information Technology for providing us an opportunity to develop this smart artifact.

REFERENCES

[1] Paul Viola and Michael J. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” IEEE Conference on Computer Vision and Pattern Recognition, 2001, pp. 511-518.

[2] Jun-Ren Ding, Chien-Lin Huang, Jin-Kun Lin, Jar-Ferr Yang, and Chung-Hsien Wu, “Magic Mirror,” IEEE International Symposium on Multimedia, December 2007, pp. 176-185.

[3] Tatiana Lashina, “Intelligent Bathroom,” Philips Research, Netherlands.

[4] Ana C. Andrés del Valle and Agata Opalach, “The Persuasive Mirror: computerized persuasion for healthy living,” Accenture Technology Labs, France.

[5] Tomoki Hayashi, Hideaki Uchiyama, Julien Pilet, and Hideo Saito, “An Augmented Reality Setup with an Omnidirectional Camera based on Multiple Object Detection,” International Conference on Pattern recognition, 2010, pp. 3171-3174.

[6] Carlos Hitoshi Morimoto. “Interactive Digital Mirror,” IEEE Proceedings on Computer Graphics and Image Processing, 2001, pp. 232-236.

[7] Hyun Hoi James Kim’ “Survey Paper : Face Detection and Face Recogntiion”.

[8] W. Zhao, R. Chellappa, A. Rosenfeld, and P. Phillips, “Face Recognition : A Literature Survey,” 2000.

[9] W.K. Teo, Liyanage C De Silva, and Prahlad Vadakkepat, “Facial Expression Detection and Recognition System,” Journal of the Institution Engineers, 2004, pp. 14-26.

[10] Hua Gu, Guangda Su, and Cheng Du, “Feature points extraction from Faces,” Tsinghua University, China.

[11] R. Brunelli and T. Poggio, “Face Recognition : features versus templates,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15, No.10, Oct 1993, pp. 1042-1052.

[12] R.S. Feris, T.E. de Campos, and R.M. Cesar Junior, “Detection and tracking of Face features in video Sequences,” Lecture Notes in Artificial Intelligence, 2000, 1793(4), pp. 129-137.

[13] Mohamed Deriche, “A Simple Face Recognition Algorithm using Eigeneyes and a Class-Dependent PCA Implementation,” International Journal of Soft Computing 3(6), 2008, pp. 438-442.

[14] M. Turk and A. Pentland, “Eigenfaces for Recogntion,” J. Cognitive Neuroscience, Vol.3, 1991, pp. 71-86.

[15] Matthew A. Turk and Alex P. Pentland, “Face Recognition using Eigen Faces,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991. pp. 586-591.

[16] M. Gargesha and S. Panchanathan, “A Hybrid Technique for Facial Feature Point Detection,” Fifth IEEE Southwest Symposium on Image Analysis and Interpretation, 2002, pp. 134-138.

[17] F. Bourel, C. C. Chibelushi, and A. A. Low, “Robust Facial Expression Recognition Using a State-Based Model of Spatially – Localised Facial Dynamics,” Proceedings of the Fifth IEEE international Conference on Automatic Face and Gesture Recogntion, 2002, pp. 106-111.

[18] Claude C. Chibelushi and Fabrice Bourel, “Facial Expression Recognition : A Brief Tutorial Overview,” http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/CHIBELUSHI1/CCC_FB_FacExprRecCVonline.pdf [retrieved : Jan, 2012]

[19] http://onionesquereality.wordpress.com/2009/02/11/face-recognition-using-eigenfaces-and-distance-classifiers-a-tutorial/ [retrieved : Jan, 2012]

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

Smart Kitchen Cabinet for Aware Home

Karuppiah Pal Amutha, Chidambaram Sethukkarasi, Raja Pitchiah
National Ubiquitous Computing Research Centre
Centre for Development of Advanced Computing
Chennai, India
{palamuthak, ctsethu, rpitchiah}@cdac.in

*Abstract***— This paper presents the design and development of a "Smart Kitchen Cabinet" which identifies the grocery items in the kitchen store. The Kitchen Cabinet is augmented with sensors to measure the weight of an item which is updated to a database whenever grocery items are placed or taken out for cooking. The jars in the kitchen cabinet are tagged with Radio-frequency identification (RFID) tag for identifying and tracking the location. The optimal placement of jars (containing different ingredients) attached with RFID tags and antennas are tested for maximum read performance and the experimental results are presented. The system also generates automated shopping list when an item reaches the defined threshold level, which is based on requirement and consumption pattern of family members.**

Keywords-Ubiquitous Computing; Pervasive Computing; Smart artifact; Smart Kitchen Cabinet; RFID.

I. INTRODUCTION

The kitchen is a very important place of a home and cooking is one of the day to day activities. The usual difficulty in a kitchen during cooking is finding the items to be out of stock.

The growing popularity of automated systems indicates the demand of the household devices to be smart and automated to support us in our daily activities. The kitchen is one ideal place where automation at various levels can be done. Daily kitchen activities include stocking kitchen cabinet in relation to necessary dietary regiment, likes, and needs, tastes etc.

Smart Kitchen Cabinet is an innovative appliance that incorporates interactive services. It is an embedded system which consists of a touch screen Liquid Crystal Display (LCD), load sensors, Radio-frequency identification (RFID) reader and tags to provide complete awareness about ingredients and availability information for better kitchen management. The features of the cabinet are: inventory management of grocery items, automatic shopping list preparation, item identification and tracking and balanced diet cooking.

The contribution of the paper is as follows: 1) a location sensing and tracking algorithm for grocery items used at home using UHF RFID and load sensors. 2) Development of an embedded inventory management for kitchen groceries. 3) Deployment results in actual kitchen environment. 4) Engineering of Smart Kitchen Cabinet. 5) Optimal placement of RFID antenna and Tags.

The remaining of the paper is organized as: Section II discusses about the related work in this area. The system architecture and overview are described in Section III. The entire system testing and field reports are presented in Section IV. The limitations of the smart kitchen cabinet are mentioned in Section V. Conclusion and future work are described in Section VI.

II. RELATED WORK

A prototype [1] called "Smart Kitchen" that enables traditional meal preparation and healthy cooking by raising awareness about facts on nutrition's present in food ingredients. The sensors are used to detect cooking activities and provide feedback to the user about nutrition information

Context-aware cooking [2] is implemented using augmented cutting board and sensor enriched knife. The cutting board is fixed with load and acceleration sensors to identify the type of food used. Microphone is used to recognize the cutting sound and a camera to identify the object.

Instrumented kitchen to automatically capture, share, and exploit semantically annotated cooking experiences has been realized in [4]. All necessary information should be observed from the user's natural course of actions during the preparation process, such that even users without any knowledge about ontologies are able to create and benefit from semantically represented recipes.

Bonanni, et al. [5] presented an augmented reality interface based on a model of the user, the task and the environment that projects information on the status of work surfaces, storage and tools directly on the objects and spaces where users direct their attention. This prototype uses range finder to measure the surface temperature of food in pans on the range. The temperature of the water in the tap is sensed using heatsink and represented by different light colors. 4D FridgeCam is an augmented reality interface that projects the contents of the refrigerator directly onto the door in such a way as to add location and time-based information. Augmented Cabinetry is an active inventory system that reduces the time required to locate items in the kitchen cabinets.

The accurate location sensing of different objects in a smart shelf using Ultra-high frequency (UHF) RFID technology is presented in [7]. Reference tags are kept in the shelf for location identification. The Received Signal

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

Strength Indicator (RSSI) value along with the tag interference level is used for locating an object.

Experimental tests for detecting the pharmaceutical items in a small cabinet using UHF RFID operating at 860MHz to 868MHz are demonstrated in [9]. They achieved a maximum of 61% full detection rate by using four antennas in a cabinet.

Our development work describes practical implementation of embedding intelligence into existing kitchen cabinet for load sensing, location tracking and automatic shopping list preparation resulting in inventory management of grocery items in the kitchen. The RFID read performance with different type of grocery items used in kitchen environment is also presented. The system achieves better read performance by using two antennas and the placement of tags as described in Section IV.

III. OVERVIEW AND ARCHITECTURE

A. System Design

This section comprises technical description of the system. RFID tags are used for identifying the item and load sensors are used for measuring the weight of items as well as locating the items in the cabinet. The RFID tags are attached to the containers and the load sensors are kept under the partitions of the cabinet. RFID antennas are mounted inside the cabinet. The placement of RFID antennas and tags [3] is tested for maximum read performance. When a container arrives/departs, the algorithm identifies the occurred event and updates in the database. The entire application is ported into an embedded platform based on Intel Atom processor. The user can interact with the system through a GUI.

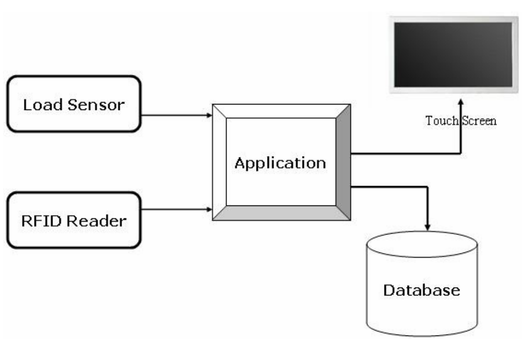


Figure 1. System Overview.

The functional representation of the system is shown in Figure 1. The sensors sense the environment and send the data to the application. The application analyzes the data, decide the action to be taken (IN/OUT) and update in the database. The user can interact with the system through the touch screen monitor. The entire system is deployed in a kitchen of C-DAC, Guest house (Figure 2) for testing purpose. State transition diagram of smart kitchen cabinet is shown in Figure 3. It shows the entire activity of the system. User has to set all necessary fields like family setting, reminder - menu setting and mapping tagID with an item. Once initial settings are done, the system is ready to function. Monitoring weight state is continuously running in the system, if any weight variation occurs, the state is transited to the finding item state which finds and detects the item.

Figure 2. Deployment of Smart Kitchen Cabinet in C-DAC, Guest House.

Detected item and its weight are stored in the database along with partition. When an item is found to be below the specified criteria level, the shopping list state is triggered. This state generates the shopping list and sent to the user through e-mail/SMS/printing according their request.

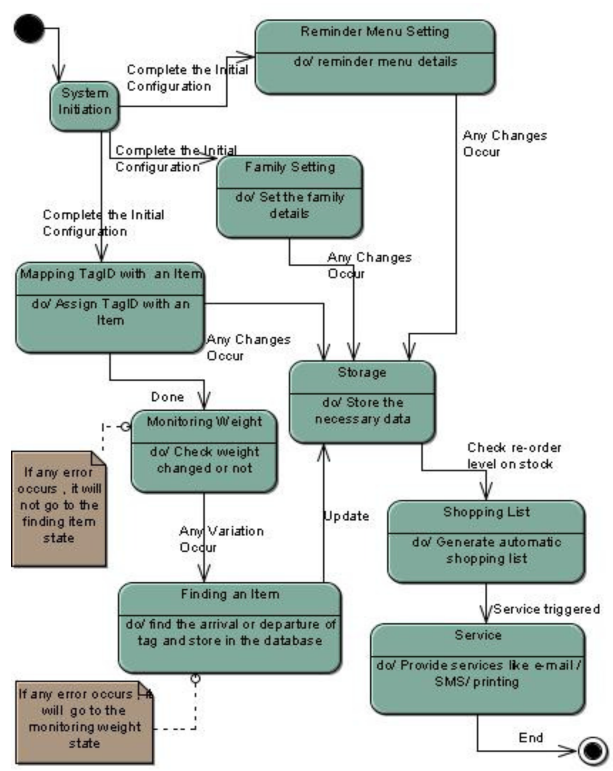


Figure 3. State Transition diagram of Smart Kitchen Cabinet

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

B. Identifying and Tracking Algorithm

The System monitors the weight over the partitions (A and B shown in Figure 2) continuously. The weight measurement is done with a minimum delay between each cycle for the load sensor to get stabilized. An event is said to occur when the difference of the weight of current cycle (ith cycle) and previous cycle (i-1th cycle) exceeds a specified threshold ($(W_i - W_{i-1}) > Th$). The occurred event may be any one of the following 1) Arrival of a new item 2) departure of an existing item 3) False Variation. If the difference value is below the threshold value, then it might be of false variation due to user's hand pressure. It represents no action has to be taken. If the weight output of the current cycle is greater than the previous weight value, then an item has arrived. The current RFID scan (T_n) is compared with the previous RFID scan (T_o) to identify the arrived tag. ($Tn \cap To$) gives the common items in current and previous RFID scan. Subtracting ($Tn \cap To$) with the current RFID scan (Tn) gives the tag ID of the new arrived item. In case of arrival event, the database is updated with the arrived tag information and weight value. Otherwise, an item has departed from the cabinet. Again, the new and old RFID scan list is compared to find the departed tag. The item is marked as "OUT" in the database. The pseudo-code of the algorithm is described in Figure 4. The available items are listed out with the details like location and quantity to the user.

//W_{i-1} Weight output of i-1th cycle.
//W_i Weight output of ith cycle.
//T_o Represents set of existing tags (Set of tags read in previous scan).
//T_n Represents set of tag read by current scan.
//Th Threshold.

```
IF((Wi - Wi-1 ) > Th) THEN
        IF(Wi  > Wi-1 ) THEN
                // A New item has been arrived
                Incoming  Item = Tn - (Tn  ∩ To )
                Item Weight  = Wi - Wi-1
        END IF
         ELSE
                // Depart of  an existing item
                Departed Item = To - (Tn  ∩ To )
                Item Weight  = 0
         END ELSE
END IF
ELSE
                // Error due to hand pressure
                //No action to   be performed
END ELSE
```

Figure 4. Pseudo-code of Identifying and Tracking Algorithm.

C. Inventory Management Functions for kitchen

Inventory management includes managing the grocery items, efficient utilization of grocery items and communication to the user [6].

1) Monthly requirement of a family:

The amount of item (Cqty) used per day is found out from the weight variation and logged in the database. From the Cqty, the required consumption (R_{qty}) of an item for a particular time period (T) is calculated using the following formula.

$$Rqty = \sum_{i=1}^{T} Cqty(i) \qquad (1)$$

2) Automatic Shopping List Preparation

$$Cq = N * Rqty \qquad (2)$$

A Minimum weight can be set for every item as a threshold limit called as critical quantity (Cq), calculated using formula (2). It denotes the quantity to be stocked for minimum number of days (N). For an example N=2 means the available quantity should serve the family for at least 2 days. When the quantity of an item goes below critical quantity; it is automatically added in the shopping list along with the quantity (Sqty) need to buy. The Sqty is calculated with the help of formula (3). Ts represents shopping interval and Aqty represents available quantity.

$$Sqty = Rqty * Ts - Aqty \qquad (3)$$

Once the shopping list is generated, it should be communicated to the user either by message/mail as per their request. The list is also displayed in the GUI.

D. User Interface

The user can interact with the system through Graphical User Interface. The following settings can be made through the GUI. User has to enter the details such as number of adults in the family , number of children in the family, number of guest (adult and children), number of days the guest will stay, alert service such as SMS or E-Mail or both and alert time.

User has to set the details like mapping tag ID with the item, brand name, empty container weight, the quantity to be stocked for minimum number of days and Shopping interval. The reminders can also be set by the user. Recipe preparation tips are provided. The system suggests a suitable recipe that can be prepared from the available groceries.

IV. TESTING OF SMART KITCHEN CABINET

A. RFID Testing

1) Materials Used: A cabinet made up of metal and wood with 4 partitions, a wooden Cabinet, UHF RFID reader from SIRIT operating at 860MHz to 868MHz, omni directional RFID antennas from Poynting, UHF passive RFID Tags operating at 860MHz to 869MHz, load sensors from loadstar, plastic containers of different height and size, porcelain containers, glass containers, Tupperware containers, stainless steel containers and grocery items are used. The cost of the materials is listed in Table I.

TABLE I. MATERIAL COST

Material	Quantity	Cost (INR)
RFID Reader	1	109031
RFID Antenna	2	11975.97
RFID Tag	20 Approx	998.00
Load Sensor	16	318560.8
Atom Processor Kit	1	70,000
Cabinet	1	25,000

2) Test 1:

The test had been carried out in the Cabinet (100 x 39 x 70 cm) made up of both metal and wood (shown in Figure 6). The test was conducted in our UBICOM laboratory.

Scenario 1:

Initially, the test was carried out with empty containers with tag attached on the side of the containers. Later the containers were loaded with grocery items. The tag detection rate of the containers was tested for different antenna position. Initially a circularly polarized antenna was mounted on the top of rack1 facing downwards. The racks were divided into grids and tested the detection rate in each grid. The read performance was around 77.7% in rack1, 20% in rack 2, 33% in rack 3 and 25% in rack4 shown in Figure 5.

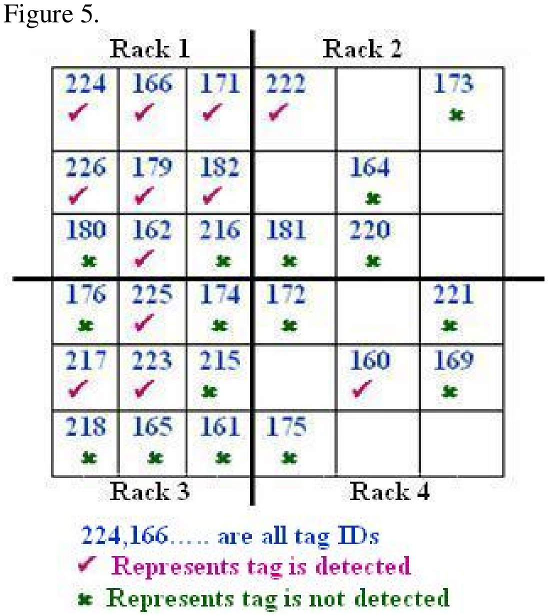


Figure 5. RFID detection result for single antenna

Figure 6. Cabinet made up of metal and wood.

a) Scenario 2:

There were several undetected tags in scenario 1 due to tag collision and tag orientation. In order to increase the detection rate, two antennas were used. One antenna was placed on the right side of the rack 2 and another on the left side of rack 3. The detection rate was 85.7% in rack 1, 100% in rack 2, 85.7% in rack 3 and 62.5% in rack 4 shown in Figure 7.

In glass and porcelain containers, the tag was read by the reader when it is fixed with some air gap as shown in Figure 8 a) and b). In stainless steel containers, the tag was not read by the reader at any position. The RFID detection rate in different containers, materials and the effect of placement of antenna and tags are listed in table. II.

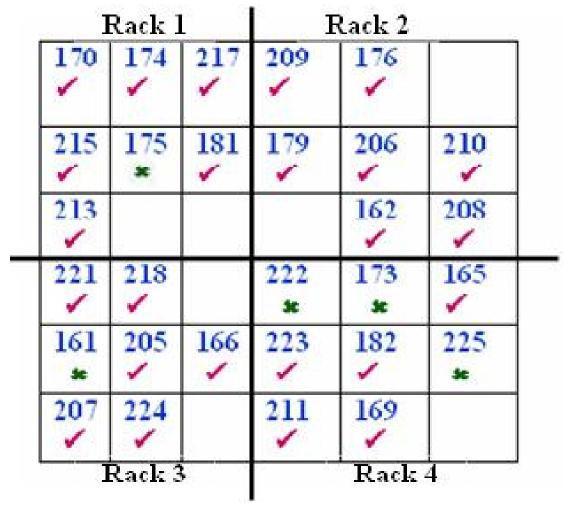


Figure 7. RFID detection result for two antennas

TABLE II. RFID READ PERFORMANCE

Container Material	Content loaded in Container	Tag Placement in Container	Antenna Placement in Cabinet	Read Performance
Stainless Steel	Anything	Anywhere	Anywhere	Very poor
Plastic	sugar, salt	Fix tag with air gap (may be in the cap, or with some space in side)	In the rack where the antenna is kept	Good
Plastic	oil	Side	In the rack where the antenna is not there	Poor
Plastic	oil	Side	In the rack where the antenna is kept	Good
Plastic	Items like dhal, rice, flour etc	Side	Any rack	Good
Small Size Plastic Containers	Anything	Side	In the rack where the antenna is	Poor

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

			not there	
Small Size Plastic Containers	Anything	Side	In the rack where the antenna is kept	Good
Glass	Anything	Side with air gap	Anywhere	Good
Porcelain	Anything	Side	Anywhere	Good

Figure 8. Tag attached with air gap in (a) Glass container and (b) Porcelain container.

b) Scenario 3:

We have found experimentally, that by adhering to the following conditions given below, the detection rate is improved to 100% (as shown in Figure 9).

- Small containers should be kept closer to the antenna.
- The sugar, salt and oil containers should be kept in the racks where the antennas are placed.
- Ensure that the containers are kept in such a way that it should not touch each other and also the corners of the rack.
- The grocery item inside the container should be below the level of the tag.

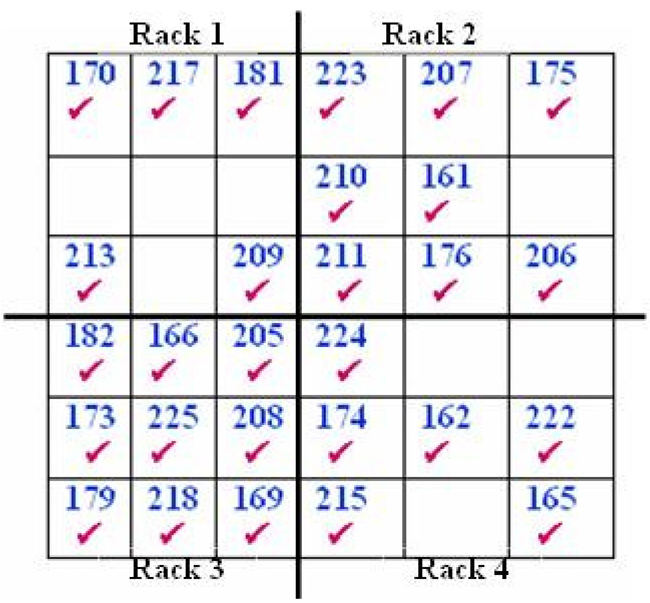


Figure 9. RFID Detection Result for scenario 3.

The RFID connection between the reader and the tags were found to be not stable in the containers having sugar, salt and oil.

3) Test 2:

The system is deployed in C-DAC guest house shown in Figure 2. The cabinet (100 x 37 x 62 cm) is made up of wood and has two partitions. The problems faced during the field testing and the solutions followed are listed below:

- Generation of shopping list:

When a container was taken out of the cabinet, the tagged grocery item was added to the shopping list. This was a logical error in the shopping list preparation module. The error was corrected by measuring the weight of the tagged grocery item after the container is placed again.

- RFID tags Arrive and Depart event problem:

Our application was based on events generated by the RFID reader. The containers with salt, sugar, and oil items, generate arrive and depart events frequently; even though the containers were kept stable in the RF field.

The tag ID and weight variation information was maintained in separate queues. When a container was kept inside the cabinet, the arrive event occurred. The arrived tag ID and the change in weight were added in the RFID queue and load sensor queue respectively. As per the queue concept, the first element of both the queues were taken out and updated in the database. Because of occurrence of false arrive and depart events, the queue concept could not be successfully used for our application.

Solutions:

The following modifications were made in our application: If both queues had an item, then the database would be updated otherwise it might be because of false arrive/depart event. Therefore, the elements were removed from both queues. Even after the modification, the problem was not solved completely.

We used the polling method for RFID tag identification. When there was a change in weight, a new RFID scan was initiated and compared with the previous scan to detect the occurrence of an event.

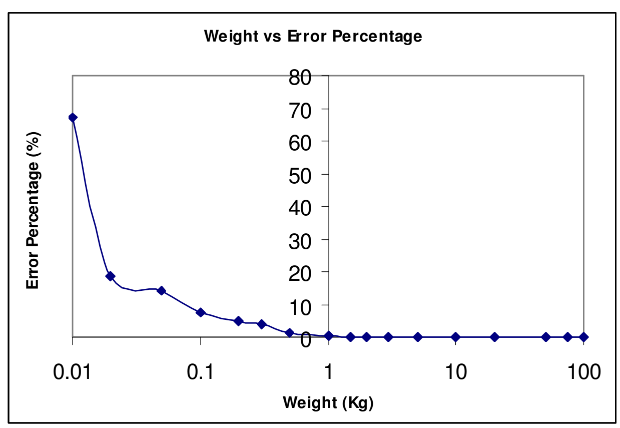


Figure 10. Weight versus Error Rate measurement.

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

- The user's hand pressure was also measured by the load sensor, while keeping the container inside the cabinet. Hence a delay was introduced in weight measurement for the load sensor to get stabilized.
- RFID connection was not stable in sugar, rice, and salt containers due to water content.
- Load sensor accuracy problem:

Solutions

Tested load sensors with standard weights. Error rate was very high for small weights and decreasing to zero approximately for the weights above 1 Kg as shown in Figure 10. It was observed that the error rate was better for the maximum load capacity of the load sensor.

V. LIMITATIONS

When an item in the container is changed, the user needs to update it. RFID connection is not stable in sugar, oil and salt containers. The placement of tags on these items can be determined experimentally for better read performance. More accurate load sensors can be used to improve the accuracy of the measurement system.

VI. CONCLUSION AND FUTURE WORK

Development of Smart Kitchen Cabinet is an effort towards kitchen automation using ubiquitous computing technologies. The system identifies the grocery items in the kitchen store. The Kitchen Cabinet is embedded with sensors to measure the weight of an item which is updated to a database whenever grocery items are placed or taken out for cooking. Based on the database information the various services offered by Kitchen Cabinet such as inventory management and automatic shopping list preparation are useful and helping us to manage the kitchen activities effectively. The optimal placement of RFID antennas and tags are analyzed for the particular cabinet and the results are presented.

The smartness of the cabinet can be further extended by adding more functionality like Nutrition-aware cooking and personalized cooking. Image processing techniques could also be explored to identify the grocery item inside the container.

ACKNOWLEDGMENT

The work is developed under National Ubiquitous Computing Research Project, funded by Department of Information Technology, Government of India. We would like to thank Department of Electronics and Information Technology for providing us an opportunity to develop this smart artifact.

REFERENCES

[1] Jen-hao Chen, Peggy Pei-yu Chi, Hao-hua Chu, Cheryl Chia-Hui Chen, and Polly Huang "A Smart Kitchen for Nutrition-Aware Cooking," Journals Pervasive Computing, IEEE, 2010, pp. 58-65.

[2] Matthias Kranz, Albrecht Schmidt, Alexis Maldonado, Radu Bogdan Rusu, Michael Beetz, Benedikt Hornler, and Gerhard Rigoll, "Context aware kitchen utilities," ACM ,Proceedings of the 1st international conference on Tangible and embedded interaction TEI '07, pp. 213-214.

[3] Rajesh Kumar Kushwaha and Ramji Gupta, "Optimization of Antenna and Tag Position for RFID Based People Management System," Proceedings of ASCNT2010, pp. 22-30.

[4] Michael Schneider, "The Semantic Cookbook: Sharing Cooking Experience in the Smart Kitchen," 3rd IET International Conference on Intelligent Environments IE 07 (2007), pp. 416-423.

[5] Leonardo Bonanni, Chia-Hsun Lee, Rob Gens, and Ted Selker, "Augmented Reality Kitchen: Task-Specific Projection in a Multi-User Work Environment," ACM UIST, 2004.

[6] http://www.invatol.com/

[7] Jae Sung Choi, Hyun Lee, Member IEEE, Daniel W.Engels, Senior Member, IEEE, and Ramez Elmasri, Member ,IEEE, "Passive UHF RFID-Based Localization Using Detection of Tag Interence on Smart Shelf," IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2011, pp. 1-9.

[8] Christoph Schonegger, Dr. Michael E. Wernle, and Dr. Burkhard Stadlmann, "Analysis of an UHF RFID System for interior position sensing," 4th European Workshop on RFID Systems and Technologies (RFID SysTech), 2008, pp. 1-5.

[9] Carla R. Medeiros, Cristina C.Serra, Carlos A. Fernades, and Jorge R.Costa, "UHF RFID Cabinet," IEEE International Symposium on Antennas and Propagation (APSURSI), 2011, pp. 1429-1432.

Copyright (c) The Government of India, 2012. Used by permission to IARIA. ISBN: 978-1-61208-225-7

Unsupervised Activity Recognition using Temporal Data Mining

Mohamed Tarik Moutacalli
Departement of Informatics
UQAC
Chicoutimi, Canada
mohamed-tarik.moutacalli@uqac.ca

Abdenour Bouzouane
Departement of Informatics
UQAC
Chicoutimi, Canada
Abdenour_Bouzouane@uqac.ca

Bruno Bouchard
Departement of Informatics
UQAC
Chicoutimi, Canada
Bruno_Bouchard@uqac.ca

***Abstract*—Excellent results have been obtained from data mining techniques in many areas. This article presents one such technique, in the context of activity recognition in a smart home. We use sequential pattern mining to analyze the history of information transmitted by the sensors, discovering thereby the frequent activities of the home occupant. Then each of the activities is temporally segmented, in order to facilitate the recognition of activities already started or even ones that are about to start. Our tests revealed that this segmentation diminished the activity search time by more than 70%, and helped predict some activities before detecting any action.**

Keywords*-*Activity Recognition; Temporal Data Mining; Temporal Segmentation; Smart Homes.

I. Introduction

The use of data mining (DM) has recently become more widespread and fashionable. The best definition of DM for us is the one suggested by Frawley in 1992 [1]: ”the nontrivial extraction of implicit, previously unknown, and potentially useful information from databases”. In particular, spatio-temporal DM [2] deals with the specific complexity of databases containing information about time or space. Temporal DM techniques are used in fields so different and so numerous that we cannot list them all: security, the stock market, electronic commerce, health, the web, etc. What is remarkable is the rarity of research which exploits these techniques in a field as large and as favorable to them as activity recognition in a smart home. The smart home, part of the recent school of thought derived from ambient intelligence [3], refers to a trend that discreetly embeds miniaturized electronic devices (sensors) in everyday objects, in order to provide real time assistance to the home occupant, based on information recently sent by the sensors or already stored in the database. The smart home can be considered as a warehouse storing a variety of data from different sensors. The large volume of data makes DM techniques the most suitable for the analysis and retrieval of knowledge, i.e., recognition of an activity (e.g., preparing meals, washing dishes, etc.), whether it has already started or is about to start.

The complexity of daily activity recognition in a smart home is due to the large number of activities that an occupant can perform. This complexity first causes a problem in creating model activities, an essential step in the process of activity recognition, where we find not only all the activities that an occupant habitually performs, but also the various actions that compose them. For example, a model of the activity preparing coffee can be composed of the actions: take cup, pour coffee, add milk and add sugar. It should also be noted that a model activity depends on the person observed, and thus it is impossible to use the same model activities for different home occupants. In our example, the activity preparing coffee for another occupant may be composed of the actions: take cup and pour coffee. For these two reasons we have included in our approach an unsupervised method for creating model activities.

The large number of activities causes a more serious problem in searching for an activity. Normally, to assist the occupant in real time, we should quickly find the required activity among all the model activities. Reducing the number of activities seems to be a natural way to speed up this search. For this purpose, we thought of using the DM technique of temporal segmentation [14]. The general idea is to create a set of time intervals for each activity, covering the periods when the occupant usually starts the activity. For example, the activity taking a shower might be segmented into two intervals: from 8:30 A.M. to 10:00 A.M. and from 7:00 P.M. to 10:00 P.M. This way, in searching for activities, we will not go through all the model activities, but only those with an interval that contains the current time. For example, if the current time is 1:00 pm, the activity taking a shower will not be considered.

Temporal segmentation proved to be a very efficient solution; moreover, it dealt with some occupant errors (due to cognitive deficiency) that no other approach has dealt with up to now. The different steps of this approach will be detailed in the next section, and the following section will be reserved for tests. Before concluding, we devote a section to related work.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

II. ACTIVITY PATTERN MINING

The first step in the process of activity recognition is the creation of model activities. The recognition agent observing the smart home occupant should have a list of all the activities that this occupant usually performs, as well as their component actions, in order to choose which activity the occupant is actually performing. Because of the large number of possibilities and the fact that model activities are unique to each person, we decided to use an unsupervised method for creating these models. In other words, the observing agent has to learn the model activities by analyzing the historical sensor data. Alireza et al. [4] have already dealt with this problem, but, as they saved all the sensor states or values all the time, activity recognition belonged more to the field of motif discovery [5]. Motif discovery is usually used in bioinformatics to speed up the detection of motifs in biosequences (e.g., DNA). So for them, an activity is a sequence that is repeated over time specifying the states or values of all sensors. Our approach, however, only saves the times and the names of sensors that have changed state or greatly changed in value, in order to reduce the size of our data warehouse without losing significant information. We interpret a small value change in a sensor as a noise and therefore ignore it. Thus, an activity is composed only of sensors that have significantly changed during the relevant time frame. Our data warehouse looks like TABLE 1:

TABLE 1. Data warehouse of active sensors per day.

Day	Sensors
1	5 2 1 4 3 2 4 6 8 3
2	6 2 1 4 7 9 2 6 1 6 9 4
3	9 8 3 7 6 1 6 9 7 1 2 1 4

For the record, the time when each sensor changed significantly has been removed from Table 1 for easy reading. For example, Day 1 is normally represented as follows: 5 (8:50) 2 (9:15) 1 (9:16) 4 (9:18) 3 (9:50) ...

As shown in Table 1, the data warehouse looks like those used in the field of market basket analysis [6], which makes sequential pattern mining [7] the most appropriate technique for detecting activities. Several algorithms have been proposed for this technique; but, to quickly find the closed frequent patterns, we chose the BIDE algorithm proposed by Wang and Han [8]. A pattern is frequent if the number of its appearances is not less than a given minimum frequency, and it is closed if it is not included in any other pattern. For example, if the activity 2 1 4 is reported, we do not report the sub-sequences such as 1 4 if they appear the same number of times as the first sequence. The BIDE idea consists of enumerating the complete set of frequent sequences, then choosing the closed ones. We start by creating a sequence tree with a root ∅. A node N at level L in the tree can be recursively extended by adding one item. Then, by removing the infrequent sequences in the sequence tree, like the node 3, the remaining nodes form a frequent sequence tree which contains the complete set of frequent sequences. Figure1 shows the frequent sequence tree built from Table 1 with a minimum frequency equal to 3:

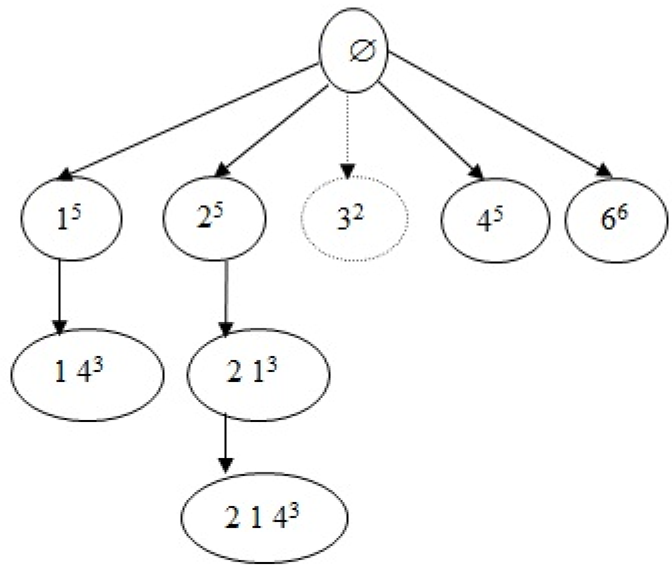


Figure 1. Sequences tree.

The customized version of this algorithm that we used allows us, at each execution, to specify the value of some very important parameters. The first indicates the minimum frequency of the activity relative to the total number of days. For example, if this value is set to 0.9 and the total number of days is ten, the activity must appear at least once in nine of the ten days. The other parameters specify the number of errors allowed for an activity and the number of errors allowed between two successive sensors. For example, if the activity is 2 1 4, then the sequence 2 6 1 4 can be interpreted as the same activity with one error that occurred between the first and the second sensor.

III. TEMPORAL SEGMENTATION

The BIDE algorithm finds all the activity models, but even allocating to each model all of its starting times, we will not have a clear and useful idea of the occupants activity habits. For this reason, we need to transform all the starting times of each model activity into time intervals which cover the periods when the occupant usually starts the activity. At any moment, then, we can use the current system time to determine the activities that he usually performs at this time, by choosing activities that have an interval containing the current time. This selection allows us to reduce the number of possibilities. Instead of searching through all the activity models for an activity started at a given time, we search only among those that the occupant usually performs at that time. Figure 2 shows the result of segmenting two activities (x and •) into two intervals each. For example, if the current time is T0, the second activity will not be selected.

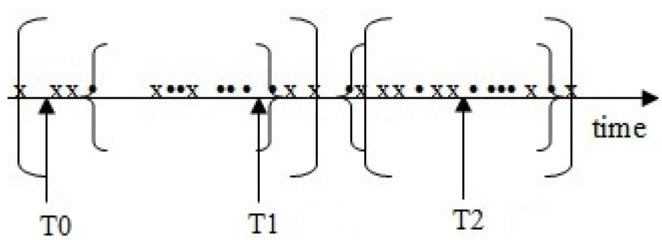


Figure 2. Example of activities temporals intervals.

The number of intervals for each activity is very important information greatly affecting the results of the segmentation. These numbers can be calculated by an algorithm such as C-means [9] for segmentation, but we found that each number is simply the maximum number of times the activity appears in a single day: it can be calculated when we visit our warehouse to get the activitys starting times.

Knowing the number of intervals into which the activity is segmented, we could use the K-means algorithm [10] where k will be equal to that number. However, as we have to sort these start times during the activity search process, we decided to develop a simpler and faster algorithm, giving the same results as the K-means. The difference between the two algorithms is that the K-means creates intervals by assigning to them the times closest to their centers, while our algorithm (see Algorithm 1 below), looping k-1 times, looks for the two successive points farthest from one another, in order to create two separate intervals.

```
Input   : a table of likely activities ActivProb ¹
Output  : a non-empty list of intervals for each
          activity in ActivProb

For each activity Activ in ActivProb
    i=1
    While i < Activ.NbrIntvl
        For each Time in Activ.TimeBeg And
        Time ∉ Table
            M = Max (NextTime - Time)
        Add IndiceM to Table
        Add 1 to i
    Sort_Asc (Table )
    Add 0 to ActivProb.Intvl
    For each T in Table
        Add T to ActivProb.Intvl
        Add T+1 to ActivProb.Intvl
    Add Activ.TimeBeg.size -1 to ActivProb.Intvl
```

1 each activity is composed of :
NbrIntvl : the number of intervals
TimeBeg : sorted table of beginning time
Intvl : table of intervals ...

Algorithm 1. Creating activities temporals intervals.

Once the intervals are created, a final check is necessary to ensure efficiency by eliminating the gaps where the activity is very infrequent and dividing any interval where two successive times are very far from each other. This last step is detailed in Algorithm 2 below.

```
Input   : ActivProb.Intvl
Output  : ActivProb.Intvl

For each Activ in ActivProb
    For each Intvl in Activ.Intvl
        If Intvl.NbrTime < (frequency/ Activ.NbrIntvl)
            Delete (Intvl )
        Else
            For each Time in Intvl
                ( NextTime – Time) > MaxDist
                Add 1 to Activ.NbrIntvl
                End = Intvl.end
                Intvl.end = Time
                Insert NewIntvl
                NewIntvl.beg = Time + 1
                NewIntvl.end = End
```

Algorithm 2. Improving results of activities temporals intervals.

IV. Activity recognition system

Temporal segmentation not only reduces the number of possibilities, but also can be used to calculate an initial priority for each activity. These priorities can tell us which activity is most likely to be initiated by the occupant, even before he acts. We have here a response to initiation errors committed by patients with Alzheimers disease that has not yet been described elsewhere in the literature, to our knowledge. Initiation errors are defined in the study by Baum et al. [11] as errors that prevent the patient from starting an activity.

The calculus of initial priorities is based on two observations:

1: The closer the current time to the end of an activity interval, the greater the likelihood of the activity. For example, going back to Figure 2 and assuming that the current time is T1; then, the activity 2 ($\bullet$) is more likely than activity 1 (x). The formula we use for this observation is: P1 = $\frac{Tc-Tb}{D}$, where Tc is the current time, Tb and D are respectively the start time and the duration of the interval.

2: The higher the frequency before the current time, within an activity interval, the greater the likelihood of the activity. In other words, the occupant was more likely to perform this activity before the current time. For example, if Tc = T2 in Figure 2, then activity 1 (x) is more likely than activity 2 ($\bullet$). The formula we use for this observation is: P2 = $\frac{Nbc}{Nt}$, where Nbc and Nt are the number of activities before the current time and the total number, within the interval.

Finally, the initial priority of the activity is: Pr = P1+P2.

Our activity recognition system assigns to each likely activity its initial priority and proposes the activity that has the greatest priority as the most likely. Once the occupant acts, we must consider his action in our decision, so these values must be updated: we add 0.5 to the priority of each likely activity involving a sensor that has just come into use, and another 0.5 if the action detected is in the right place. For example, if we have two activities: 2 1 4 and 4 2 5 7, and if we detect that the first sensor that has been activated is the sensor 4, then the priority of the first activity will be augmented by 0.5 and the second by 1.0. After each action the priority of each likely activity is updated and the activity with the greatest priority is always proposed as the most likely.

V. Validation

To test our approach, we used the database in van Kasteren et al. [12]. These data were recorded by observing an occupant for 28 days, using 14 sensors which detected seven distinct activities: 1: 'leave house'; 2: 'use toilet'; 3: 'take shower '; 4: 'go to bed '; 5: 'prepare breakfast '; 6: 'prepare dinner '; 7: 'get drink'.

Concerning the creation of model activities, the results of our tests were a little different. In fact, just five of the seven activities were discovered. The first reason for this difference is that the occupant was used to perform the activity leave house just after activity 3, take shower, which meant that our algorithm recognized the two activities as one. This difference does not really matter because the occupant will be assisted in both activities and the next action can be predicted, but the use of other temporal information, such as the average duration of an activity or the maximum time between two successive sensors, may make such assistance more efficient. The fourth activity, going to bed, was not detected because it is carried out around midnight, which usually divides its actions between two days. This is a problem, not just for BIDE, but for all similar algorithms. In fact, BIDE created more than 71% of the model activities, but it assisted the occupant in more than 85% of his activities.

To test the temporal segmentation, we used just twenty days for the segmentation and we devoted the other days to tests. Results were more than satisfactory. As shown in Figure 3, which shows the activity intervals, the number of likely activities was reduced from 30 to 70% depending on the time of day.

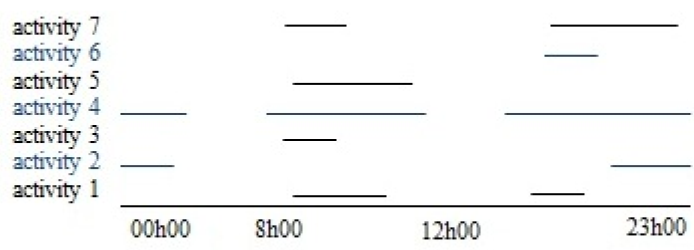


Figure 3. Representation of activities temporals intervals.

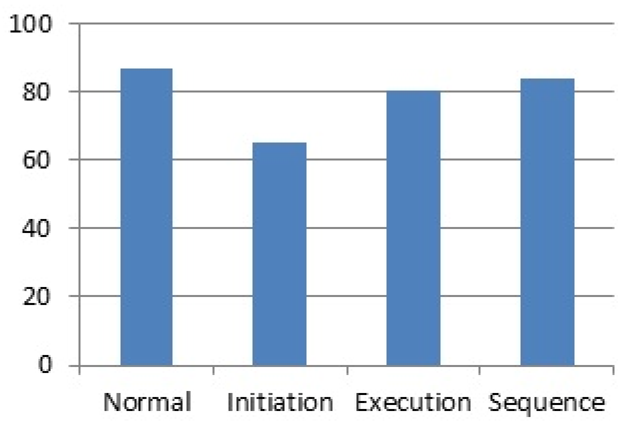
It should be noted that the number of activities that a person normally performs daily is well above the seven activities that were used in these experiments. This would necessarily increase the percentage of reduction of the number of likely activities. However, we must note another important problem in this segmentation. The perception of a single day as a time line gives the result that two times very close to each other seem to be very far. For example, 11:59 p.m. and 0:01 a.m. are considered to be almost 24 hours, not two minutes, apart. If we need to create an interval, then it will be 0:01 until 11:59 p.m., because the start time is always less than the end time, while the interval should be from 11:59 p.m. until 0:01 a.m. The difference between the two is enormous and will greatly affect our segmentation.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

That is why the final step of segmentation, the audit, is very important because it will divide such an interval which took almost all day into two small intervals, giving a very similar result to that intended.

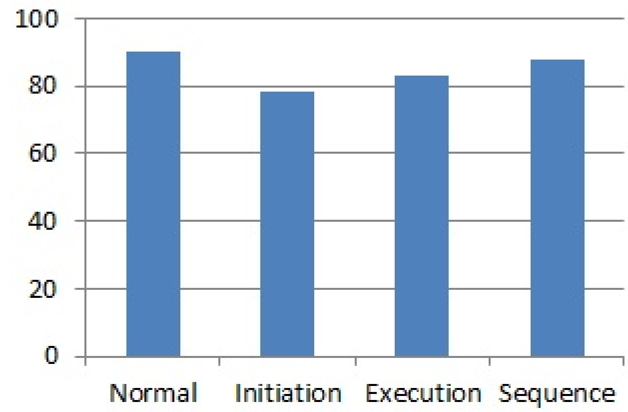
For the occupant of the smart home, who in our case is a person with Alzheimer's disease, we have introduced some errors in the tests to simulate errors likely to be committed by the occupant, which are defined by Baum [11] as:
Execution errors: the patient forgets or adds actions that have nothing to do with the activity.
Sequence errors: the patient performs in a disorderly manner the various stages of the activity.

Graph 1 shows the percentage of detected activities. The column names have the following meanings:
- Normal: all the actions that compose the activity were done.
- Initiation: no action has been done.
- Execution: some irrelevant actions were added or deleted.
- Sequence: actions were not done in order.



Graph 1. Percentage of detected activities.

In general, the results were very satisfactory. In the first column, the activities that were not detected were started at an unusual time by the occupant. We are sure that the number of these activities would decrease if we observe the patient for a longer period. The third and fourth columns show that our system has responded well to execution and sequence errors. The decrease in activities recognized, in comparison with the first column, is explained by the similarity of certain activities with respect to starting times and component actions. In the second column, the relatively low percentage of activities recognized is mainly due to periods when the segmentation has reduced the number of likely activities by only 30%. The results of the second column are improved when we remove the interval of an activity that just has been recognized, so that it would not be considered again in the same period. Graph 2 presents the new results obtained.



Graph 2. Percentage of detected activities after removing the interval of an activity that just has been recognized.

VI. Related Work

Our approach explores the new possibilities offered by the emergence of techniques of data mining and exploits them in the field of activity recognition. It is mainly based on temporal segmentation, which plays a role in the feasibility of this approach and greatly affects the results. Temporal segmentation has been used in various studies of activity recognition, but never, to our knowledge, as described in this paper. In Spriggs et al. [14], for example, cameras are used to observe the patient. Then temporal segmentation divides the movement of the observed patient into actions, in order to understand and create movement patterns that will help recognize the activity. In other words, they try to break an activity into time intervals, where each interval represents an action. These actions will therefore consist of several frames of a video sequence, and they will be classified in order to create model movements that will be used for the future detection of activities. The results of this approach, according to Spriggs et al., were difficult to assess, but we can imagine the heavy processing of video and its impact on an application that needs to respond in real time. In addition, the use of cameras is still subject to debate because it does not preserve the privacy of the occupant.

Harvey et al. [15] have used segmentation to analyze the history of the work of developers. They see time as a sequence of points. Segmentation is to combine several consecutive points in order to create a segment. The goal is to reduce the number of points to a smaller number of segments. The only problem with this technique is that the segments do not perfectly represent the input data. The segments are created by trying to minimize errors, while in our approach the time intervals summarize all the input data.

Jakkula and Cook [13] used time to find relations between events. They defined 13 relations, like event 1 after

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

event 2 or event 1 finished event 2, etc. These relations helped them to predict an event when the usually previous event happened. They were then able to make many other interesting deductions, like: an event must end because it is followed by another event that just finished. The problem with Jakkula and Cooks work is that they apply their algorithm directly on the data warehouse, so that all deductions are about events and not activities. As an event may be included in many activities, we will have many relations for the same event, and we will be confused as to which relation we should apply. We still think that this algorithm is very interesting and may improve our activity recognition system. In fact, we could use it after our first step of activity pattern mining, when the relations are between activities and not events. After that, they may be converted as probabilities or priorities that can help calculate the most likely activity in our activity recognition system.

VII. Conclusion and future works

In this study, we noticed the huge benefits of the use of data mining techniques in the area of activity recognition in smart homes for cognitive assistance to Alzheimer's patients. We have seen how these techniques were able to respond to different problems in this area. First we used them to create unsupervised and personalized model activities. Then they helped us reduce the number of likely activities by using the current time to eliminate the unlikely ones. They also allowed us to solve the problem of initiation errors of Alzheimer's patients by suggesting an activity to the occupant if the current time exceeds the end of an activitys interval without the system detecting that the activity was performed. Finally, the complex problem of intersecting activities can be solved: if the occupant is used to perform several activities at once, upon detecting an activity, all the intersecting activities will be detected as a single activity, and our approach will therefore be able to assist him.

However, the success of this approach depends on a good temporal segmentation. If, for example, the occupant tries to perform an activity at a time outside its intervals, this approach will be unable to recognize the activity even if it is contained in the model activities.

Several improvements can be made to this approach, such as trying to have an average duration of the activity in order to recognize if the occupant has problems finishing it. We may even have a maximum time between two successive sensors in order to know when to trigger aid to the occupant. These improvements and others will be included in our next article.

References

[1] J. William Frawley, G. Piatetsky-Shapiro, and J. Christopher Matheus, "Knowledge Discovery in Databases: An Overview", Ai Magazine, 1992.

[2] F. John Roddick and M. Spiliopoulou, "A bibliography of temporal, spatial and spatio-temporal data mining research", ACM SIGKDD Explorations Newsletter, Volume 1 Issue 1, June 1999.

[3] Fariba Sadri, "Ambient intelligence: A survey", ACM Computing Surveys (CSUR), Volume 43 Issue 4, October 2011.

[4] A. Vahdatpour, N. Amini, and M. Sarrafzadeh, "Toward Unsupervised Activity Discovery Using Multi Motif Dimensional Detection in Time Series", IJCAI'09 Proceedings of the 21st international jont conference on Artifical intelligence, 2009.

[5] D. Yankov, E. Keogh, J. Medina, B. Chiu, and V. Zordan, "Detecting time series motifs under uniform scaling", KDD '07 Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 2007.

[6] R. Michael Wick and J. Paul Wagner, "Using market basket analysis to integrate and motivate topics in discrete structures", SIGCSE '06 Proceedings of the 37th SIGCSE technical symposium on Computer science education, 2006.

[7] R. Nizar Mabroukeh and C. I. Ezeife, "A taxonomy of sequential pattern mining algorithms", ACM Computing Surveys (CSUR), Volume 43 Issue 1, November 2010 .

[8] J. Wang and J. Han, "BIDE: Eficient Mining of Frequent Closed Sequences", 2007.

[9] S R. Kannan and R. Pandiyarajan, "New modified fuzzy C-means for determination of proper structure in dataset", 2009.

[10] A. Vattani, "k-means requires exponentially many iterations even in the plane", 2009.

[11] C. Baum, D. F. Edwards, and N.Morrow Howell, "Identification and Measurement of productive behaviors in senile dementia of the Alzheimer type", 1993.

[12] T. van Kasteren, A. Noulas, G. Englebienne and B. Krse, "Accurate Activity Recognition in has Home Setting", 2008.

[13] R. Vikramaditya Jakkula and J. Diane Cook, "Mining Sensor Data in Smart Environment for Temporal Activity Prediction", 2007.

[14] H. Ekaterina Spriggs, F. de la Torre, and M. Heber, "Temporal Segmentation and Activity Classification from First-person Sensing", IEEE Workshop on Egocentric Vision, CVPR 2009, June 2009.

[15] H. Siy, P. Chundi, and M. Subramaniam, "Summarizing Developer Work History Using Time Series Segmentation", May 2008.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

Smart House with Smart Electricity

The need for a DC Bus

Yuval Beck
Faculty of Engineering
Holon Institute of Technology- HIT
Holon, Israel
e-mail: beck@hit.ac.il

Gadi Golan
President
Holon Institute of Technology-HIT
Holon, Israel
e-mail: golan@hit.ac.il

Edna Pasher
Edna Pasher PhD & Associates
Tel-Aviv, Israel
e-mail: edna@pasher.co.il

***Abstract*— The following work describes an ongoing idea that becomes more logical in the future of domestic and military electricity. This paper introduces the necessity for using a DC bus system in private homes as well as introducing an innovative bi-directional inverter for making the energy flow to and from the grid. The bi-directional inverter also enables smart control of the energy flow with harmonics cancelation, and the ability to be adjusted to the future demand of smart grids and multiple tariffs. The research also includes social aspects as an integrated subject in the control schemes of the electrical devices.**

Keywords- Smart house, Inverter, Smart grid, control circuits.

I. Introduction

THERE is a lot of interest in recent years in Smart Grid application. Dealing with these smart grid new concepts brings us back to basics and raises the question whether the current way of providing and consuming electricity is the correct or the best available way [1]. As for today, most of the electrical appliances that can be found in a typical house work on Direct Current (DC). A computer is assembled of DC circuits, TV is based on DC circuits, lighting and heating can also work on DC. The only appliances which require Alternating Current (AC) are the air-conditioners and the rotating machine based appliances, such as a washing machine and dryer. Although these appliances consume heavy loads, they consume only 30-60% of the total load. On the other hand, transporting energy for long distances is known to be more efficient in AC. Another factor that needs to be added to the new equation of the future energy world is the addition of renewable energy as a power source. Photovoltaic energy is DC and wind power which is AC based power sources. There is work done on DC distribution systems based on Photovoltaic generation and the electrical control strategy [3] , as well as on DC distributed systems for residential applications with the use of Bi directional Inverter [4].

The above mentioned situations lead to the question whether AC power in our home is the right way of providing our loads with the power [4],[5]. Moreover, the assimilation of communication and control capabilities in the future Smart grid enables us to control the efficiency of the power flow according to various criteria. These criteria can be influenced by the status of the total grid. As an example, let us consider a grid with over 30% renewable energy, and for stability issues the grid has to control the power flow from all renewable sources, namely, the photovoltaic inverters will not necessarily work in Maximum Power Point (MPP).

This paper reviews the idea of combining a local DC bus grid in a few applications for increasing the total efficiency of energy transformations in the overall grid. A Bi-Directional Inverter (BDI) for enabling the dual DC/AC and AC/DC power flow with reduction of harmonics and control is presented. The converting systems are treated as Time Variable Transformers with advanced control that combines also non-electrical parameters. The paper also deals with micro-grids and synchronization issues by using the DC bus and the social aspects of controlling this electricity in the Smart Grid environment.

II. Domestic DC Bus

In recent years private installations of PV generators and wind generators are becoming common in many countries. There are differences in the supporting laws in each country but in general these renewable generators are a financial business in which the investor wants to restore the investment on the system and make a profit on the electricity he sells to the grid. For achieving this goal, the current installed inverters are equipped with Maximum Power Point Trackers (MPPT) for maximizing the power extracted to the grid.

When renewable energy will reach a dominant percentage of the grid current, other issues of stability of the grid will rise. In this case not all renewable generation will be allowed to penetrate the grid at all times. Furthermore, future storage that will be available, even in domestic facilities, will require a more complex control on the power flow.

The fact that most of the appliances in our homes are operating on DC raises the question whether it is logical to have a PV generator on the roof, then invert the DC to AC, then convert the AC back to DC to power all of our electronic circuits. This two conversion process is obviously not efficient since each energy conversion reduces the total efficiency of the process.

It is probably more logical to have a separate DC bus at our home, parallel to the AC grid (that will power all the AC appliances such as air conditioners etc.) as seen in Fig.1. In this case the future computer, TV and all other DC appliances will not need to have an AC/DC rectifier in their

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

input power supply. This yields to more efficient and cheaper electrical appliances.

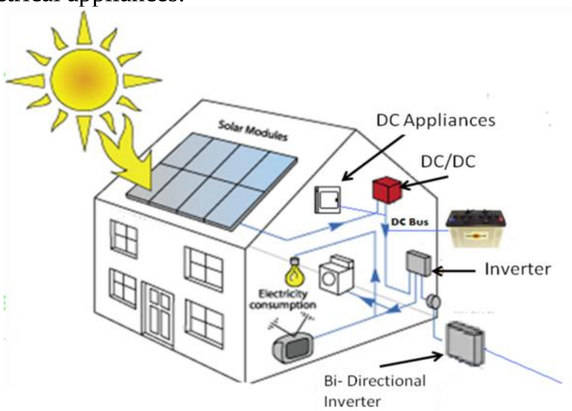


Figure 1. A combined DC and AC grid in a house

Fig. 1 shows the components that are necessary for assembling an integrated DC and AC home.

The house has two power grids: one is an AC grid and the other is the DC bus. The PV panels on the roof are connected to a DC/DC converter that has an MPPT and control schemes for various grid constrains algorithms. This converter powers the DC bus. Moreover, there are storage capabilities that are connected to the DC bus also. Furthermore, there is an option of charging the DC bus from the grid with a bidirectional inverter that will be discussed later on. The purpose of the bi-directional inverter is to enable bidirectional power flow in cases such as:

- PV energy must be controlled due to grid constrains.
- PV is not available and DC power is required.
- Grid electricity is cheaper and DC power is required for charging the storage device.

A. The bidirectional inverter

The core of the system is the bidirectional inverter. A schematic drawing of the whole system is presented in Fig. 2.

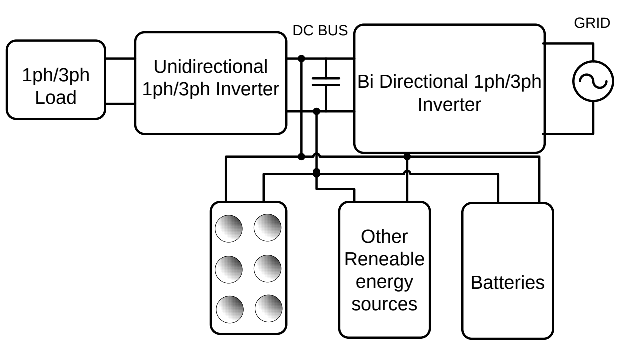


Figure 2. A block diagram of the electrical system in a future house.

In Fig. 2, it is clear that the gateway from the domestic electric grid to the "main" grid, namely, the utility, is the bi-directional inverter. This inverter, in one direction, can transfer power flow from the power grid towards the loads and DC bus. In the other direction, it can transfer the power from the house back to the grid. The advantage of this device is the ability to control the power flow as well as controlling the power quality. The bidirectional inverter is equipped with a bidirectional power factor corrector and control that is also influenced by dynamic tariffs, social aspect such as nonintrusive algorithms and electrical efficiency control.

This system is designated for small to medium loads of up to 15kW with modularity capabilities for expansion. The system is connected in one side via a Bi Directional one phase or three phase inverter to the grid. On the DC side of the inverter there is a unidirectional one phase or three phase inverter for supplying the load (home, commercial, industrial).

When no renewable energy is available, the system is acting as an active filter for reducing harmonics [6]. When renewable energy is available, it produces energy directly to the DC bus and will smoothly supply energy. The system will then divide the amount of energy that the grid supplies and the energy from the renewable sources. If there is excess supply from the renewable energy sources, the system will return energy back to the grid or store it in the storage bank. Moreover, if at any time electricity prices decrease, the system can store energy in the storage bank for later use.

The main challenge is the development of optimal control strategies for energy usage aiming to maximize customer economic savings, while preserving user preferences and quality of experience, based on a proper subscribed power supply contract. The key aspect of the research is the optimal trade-off between the exploitation of convenient fares and the overload management, by considering the following electric devices:

- Loads characterized by different controllability degrees.
- Storage units.
- Uncontrollable Micro Generation.
- Other small size traditional power plants.

Preliminary evaluation of the problem suggests that the control signals to be optimally generated consist of:

- The best times to run programmable loads.
- Power recharge/release evolution of the storage units.
- Evolution of generation set-points of traditional power plants.

All these aspects need to be taken into account, together with time varying energy prices, estimated power consumption from not programmable loads and the forecast of generation from renewable and uncontrollable units. The solution of integrating distributed control algorithms with dynamic programming approaches will be investigated. The Implications on the Power Grid are considered too.

The state of the art technology offers inverters with harmonics reduction. The common method is to use simple PWM controlled inverters and multilevel topologies for reduction of the harmonics content. Other techniques involve a controller that samples the current and phase shifts the current for achieving PFC. This research is offering a different approach by namely considering the Inverter/Converter as a Time Variable Transformer (TVT)

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

whose power stage is designed to be a simple Inverter and with flexibility and enhanced characteristics which are obtained by advanced control schemes. In this research the following control schemes are being developed, analyzed, simulated and implemented:

- Power flow control that enables the BDI to perform as an Inverter or as a Rectifier by demand.
- Control scheme to reduce harmonics at inversion.
- Control scheme to reduce harmonics at rectification.

In general, the system is considered as a TVT, as seen in Fig. 3,[7],[8]:

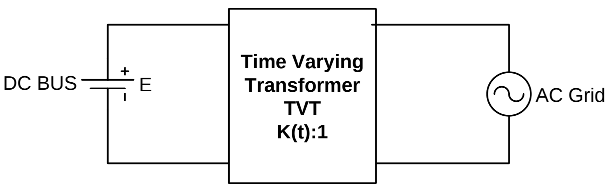


Figure 3. TVT presentation of Bi- Directional Inverter

The DC side is held at voltage E by either a storage bank or a large capacitor. The AC grid side is either a one phase or a three phase system. Namely, for the one phase system, the voltage is:

$$v(t) = V_m \sin(\omega t + \theta) \quad (1)$$

For the 3 phase system the AC grid will have three phase voltages as follows:

$$\begin{aligned} v_1(t) &= V_m \sin(\omega t + \theta) \\ v_2(t) &= V_m \sin(\omega t + \theta - 120^0) \\ v_2(t) &= V_m \sin(\omega t + \theta + 120^0) \end{aligned} \quad (2)$$

In the case of a three phase system, the power injected or drawn from the grid is constant and in the case of a one phase system, it is time dependant and fluctuating in a frequency that is double the grid's frequency [9]. By applying an appropriate control signal k(t), the AC and DC sides can be matched in a way that harmonics will be eliminated. For example, in an AC to DC conversion in a one phase system, the control signal for eliminating the harmonics, the transfer ratio of the TVT is: (when considering a realistic system with an internal resistance of the DC source of r) [8], [10]:

$$k(t) = \frac{2V_m \sin \omega t}{E\left[1 + \sqrt{1 + \frac{2rI_m\left[\cos\phi - \cos\left(2\omega t - \phi\right)\right]}{E^2}}\right]} \quad (3)$$

In a three phase system, the control due to the constant power behavior is simpler. However, the hardware is more complex and therefore there are more control signals (for each phase). The control can be represented as:

$$\left[k(t)\right] = \begin{bmatrix} k_1(t) & 0 & 0 \\ 0 & k_2(t) & 0 \\ 0 & 0 & k_3(t) \end{bmatrix} \quad (4)$$

Each element in the matrix (4) can be subdivided into the signals relative to each one of the switches (FET, IGBT, etc). In this way each element can be represented as a sub matrix.

B. *Local and Grid Control*

The existence of a BDI at the entrance to each house yields a lot of opportunities. Since each BDI is equipped with a unit that has a control and ability to be programmed, the general system can be divided into two main controls, as seen in Fig. 4.

In Fig.4 the two control schemes are presented. Each BDI is controlled for maximizing the energy efficiency in the house. This control has distributed characteristics. On the other end, the units are connected to the smart grid and have a centralized grid control. This control can also be in a few hierarchies (neighborhood, city, country, etc.). The grid control takes into account constrains of the grid, namely, the ability of the grid to absorb energy from renewable energy sources, reducing the losses on main transmission lines, overload and frequency changes in the grid, stability etc. The local domestic control then responds to the grids demands and can change the profile of the prosumer for achieving efficiency in all levels.

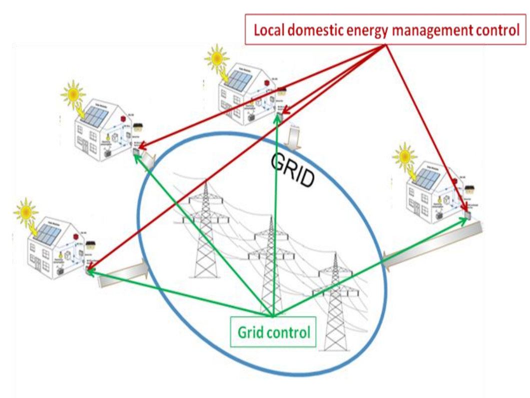


Figure 4. A block diagram of the electrical system in the future home.

III. MICRO-GRID APPLICATION

Another important application can be micro-grids, as seen in Fig.5.

In this example a small grid is being established (such as in field deployment applications). The grid is constructed by small generators, storage devices, renewable (PV and wind), DC and AC loads and an optional grid connection. The mixture of sources types and load types makes a single AC grid very difficult to implement.

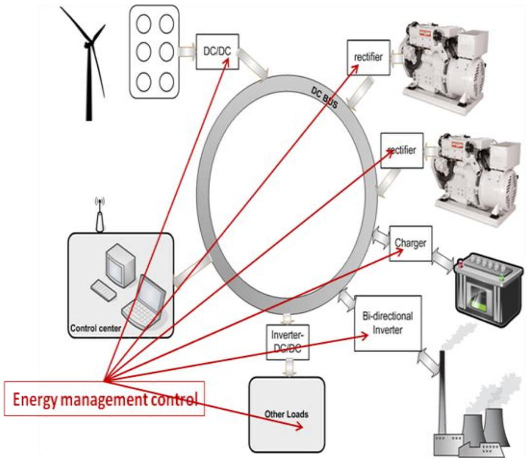


Figure 5. A DC bus in a micro-grid application

In a micro-grid, the synchronization of the various sources is difficult due to stability issues. The DC bus is then the natural solution due to the fact that no synchronization is necessary. All sources are contributing the power to the common DC bus via a DC/DC converter or an AC/DC inverter. The optional BDI enables connection to the general grid for emergency power need or for injecting power to the grid.

IV. SOCIAL ASPECTS OF SMART ELECTRICITY

This research will engage all stake holders up-front. The assumption is that in order to speed up the adoption of any new technology, all stake holders must be involved [11]. In the case of the DC Bus, the stake holders have to be mapped. First, they have to be identified and then discover their real needs. Another aspect that is highlighted here is the fact that in a large project a full engagement of the end-users is necessary. This methodology is named: 'user centered design' (UCD) or a 'living laboratory'. Another important aspect for knowledge-based development is the notion that change can happen only if all stakeholders are involved A Living Lab has to be created in which the consumers will become full partners in the iterative research and development cycle. This methodology was successfully used in other research projects where developing systems with state of the art technology of wearable computing was introduced into different environments [12], (www.wearitatwork.com). Social aspects will thus be studied using an Action Research approach producing guidelines to the developers to be followed by further ethnographic research and interviews for data collection, once prototypes have been introduced and then followed by further guidelines to the developers of the new technology. The conclusions of the social research will be implemented as control schemes for controlling the electrical equipment as non- intrusive operation.

Thus, this research is interdisciplinary with better probability of quick exploitation by all stakeholders.

V. CONCLUSION

This paper describes an ongoing research which aims to enable the insertion of a separate DC bus to private homes and other applications where it is natural to use this idea. The main reason for utilizing this idea is the fact that most loads are DC operated as opposed to the load characteristics in the past (30 to 40 years ago). The system is based on BDI with control schemes that are designed for maximizing the power efficiency, minimize the harmonics for maintaining power quality and take into consideration the social aspects of the prosumer.

REFERENCES

[1] M. Hashmi, S. Hänninen, and K. Mäki, "Survey of Smart Grid Concepts, Architectures,and Technological Demonstrations Worldwide", 2011 IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America), pp. 1-7, Nov, 2011.

[2] L. Zhang, K. Sun, Y. Xing, L. Feng and H. Ge, " A Modular Grid-Connected Photovoltaic Generation System Based on DC Bus", IEEE Trans. On Power Electronics, Vol. 26, No. 2, pp. 523-531, Feb, 2011.

[3] T.-F. Wu, Y.-K. Chen, G. –R. Yu and Y. –C. Chang, "Design and Development of DC- Distributed System with Grid Connected for Residential Applications", 8th International Conference on Power Electronics-ECCE, Asia, pp. 235-241, May 30th- June 3rd, Korea, 2011.

[4] D. J. Hammerstrom, "AC Versus DC Distribution Systems --- Did We Get Right?," IEEE Power Engineering Society General Meeting pp. 1-5, 2007.

[5] A. Mohamed, , and O. Mohammed, "Smart Power Flow Control in DC Distribution Systems Involving Sustainable Energy Sources", 2010 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America pp.372-379, 8-10 Nov. 2010

[6] M.H. Rashid, Power Electronics Handbook, third edition, Academic Press, 2001.

[7] S. Singer, " Power Conversion and Control with Zero AC Current Harmonics by Means of a Time-Variable Transformer", IEE Procidings, Vol. 131, PT. G, NO. 4, pp. 147-150, AUG. 1984

[8] S. Singer, " Modeling and Synthesis of Harmonic-Free Power Conversion and var Compensation Systems", IEEE Trans. On Industrial Electronics, Vol. IE-34, No.2, pp. 257-161, May, 1987.

[9] E. Paal, Z. Tatai, "Grid Connected Inverters Influence on Power Quality of Smart Grid" , 14th International Power Electronics and Motion Control Conference, EPE-PEMC 2010, pp.35-39, 2010.

[10] I.A. Khan, R.W. Erickson, "Synthesis and Analysis of Harmonic-Free Three-Phase Inverters", IEEE Transactions on Power Electronics, Vol. 9, No. 6. pp, 567-575, Nov. 1994.

[11] M. Armstrong, D.J. Atkinson, C. M. Johnson, and T.D. Abeyasekera," Low Order Harmonic Cancellation in a Grid Connected Multiple Inverter System Via Current Control Parameter Randomization", IEEE Trans. on Power Electronics, Vol. 20, No. 4, July 2005.

[12] E. Pasher, Z. Popper and H. Raz," WearIT@work: a wearable computing solution for knowledge-based development", *Int. J. Knowledge-Based Development*, Vol. 1, No. 4, pp. 346-360, 2011.

MyUI Individualization Patterns for Accessible and Adaptive User Interfaces

Matthias Peißner, Doris Janssen, Thomas Sellner
Competence Center Human-Computer Interaction
Fraunhofer IAO
Stuttgart, Germany
{matthias.peissner, doris.janssen, thomas.sellner}@iao.fraunhofer.de

***Abstract*—User groups for interfaces to interact with smart environments are highly heterogeneous. Furthermore, such user interfaces often run on different devices, ranging from smartphones to interactive TVs and others. Dynamic user interfaces, which adapt to different devices and diverse user needs at the same time in order to provide a universal access have received little attention so far. Our MyUI system follows a design pattern based approach to adaptive and accessible user interfaces. In this paper, we present the MyUI individualization patterns. They build on information about the user, the context and the devices, which is gathered through system interaction and sensors. The individualization patterns determine global characteristics of the user interface to adapt interaction mechanisms and presentation formats to achieve an optimal fit for the user's abilities, situation and equipment. Through this work, an understandable and human-readable way for editing accessibility rules for a system is achieved.**

Keywords - Design Patterns; adaptive systems; accessibility; user characteristics

I. INTRODUCTION

Modern sensor and agent technologies in smart homes aim at interpreting the user's behavior and intention as well as contextual information of the environment and the used devices. Through these pervasive technologies new appliances for elderly or handicapped persons due to physical/cognitive restriction can be developed for smart environments. For example, a smart TV offers several services such as weather, email or cognitive and physical exercises [6]. Especially, a user group of people with certain disabilities needs a lightweight and unrestricted access to all smart multimedia devices and services through an individual interface [10]. Therefore, it is necessary that these smart multimedia interfaces are self-learning and self-adaptive to offer great opportunities for accessibility and improved usability for a broad range of diverse users and devices.

In this paper, we present the MyUI individualization patterns for adaptive user interfaces. Within the MyUI project, an infrastructure is defined for accessible user interfaces in smart environments, which self-adapt during runtime in order to cover individual needs and limitations. We will show how this infrastructure collects user, device and context information and generates within an ongoing process a current user profile, device profile and user interface profile. The MyUI adaptation relies on a repository of user interface design patterns, which are selected and put together according to the information within the current profiles. A major objective is to reduce the need for configuration or user enrolment.

In this paper, we will set the focus on the MyUI individualization patterns, which define rules for transforming the collected data from the infrastructure to an individual and adapted user interface. We show related work, then explain the concept and the pattern template, and we will show two detailed examples of such interaction patterns.

II. RELATED WORK

For the personalization of user interfaces, mainly two approaches exist. Adaptable user interfaces provide the user with mechanisms to actively customize the interface, whereas adaptive user interfaces initiate and perform dynamic adjustments autonomously ([4][7]).

Both approaches have their pros and cons. The most important advantage of adaptable systems is that the users are in total control of the individual appearance of their user interface. This supports the understandability and traceability of modifications from the user's perspective ([5]). However, this advantage is at the same the main shortcoming of adaptable user interfaces – especially when personalization aims at increasing the accessibility. Users with disabilities and lower levels of ICT (information and communication technology) literacy would benefit most from personalized user interfaces. They often have severe problems with standard configurations. Customization dialogues, however, are a significant barrier – even for abled users [8]. For that reason the system shall provide both functionalities of adaptable and adaptive systems. A good approach for this has been developed within the mixed-initiative system MICA, where the user gets system generated recommendations and decides, if he wants to accept them [3].

But, if the system should be more proactive (e.g., when it comes to adapt a user interface to reach a higher accessibility), it requires understanding of the application state, the current environment and user context as well as the specification of the current device at runtime. Blumendorf et al. [1] describe an approach how to combine models (user, device etc.) and the current application state in order to build adaptive user interfaces. Therefore, three abstract executable types of models – definition elements, situation elements and executable elements – were introduced. These elements

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

contain static and dynamic variables, which are set at application start or dynamically during runtime. The utilization of these models makes the system aware of context information as well as knowledge about the current application state. Through the connection between context models and application state, the system can dynamically adapt to changes in the environment or user behaviors. However, the knowledge about generating user interfaces is hidden within the system and not visible for developers.

The conceptual basis of the MyUI adaptation Framework is a design pattern repository that includes proven design solutions for optimal usability and accessibility. In [9] we defined requirements on adaptive user interfaces to improve accessibility and describe a modular approach to user interface development, which relies on the combination of specific user interface design patterns. Borchers [2] suggested a structure of design patterns. This structure is used for the MyUI patterns. The patterns are related to other patterns of different types and levels of abstraction and linked with software components that can be extended for developing adaptive user interfaces. The MyUI design pattern repository [12] is publically available for reviewing, refining and contributing by other experts in the field.

III. Parameters Influencing the Adaptation of Accessible User Interfaces

For the User Interface generating process it is necessary to gather environment, user and device information. This information is collected by several external and internal sensors, which are installed in the user's smart environment and within the used devices. The MyUI Context management system interprets and updates all this information and categorized them in two characteristics: User characteristics and device characteristics. The transformation of the raw sensor data follows a specific ontology, which follows the Open Ambient Assisted Living (AAL) framework [11].

A. User Characteristics

User Characteristics give information about the user, his abilities, disabilities, preferences, and about his current environment (see Table I). All these values are stored in a User Profile, which is continuously updated by the MyUI context manager. The value of the user profile is specified as a numeric value between zero and four. For instance, if a user has a visual acuity of 0.5 that means he has a nearly perfect visual ability. In contrast, a user with a value of 3.8 in hearing is nearly deaf. These values are defined within the MyUI pattern repository.

TABLE I. MYUI USER PROFILE VARIABLES

Name of variable	Description
Visual Acuity	Ability to perceive what is displayed on the screen
Field of Vision	Describes how limited the field of vision of the given user is.
Ambient Light	The amount of ambient light at the users place.
Ambient Noise	The amount of ambient noise at the users place.
Hearing	Describes how limited the user's ability to hear sounds is.
Language Reception	Ability to understand written or spoken language
Language Production	Ability to speak and write language
Understanding Abstract Signs	Ability to understand abstract signs and pictograms
Attention	Ability to handle multiple things at the same time, resp. focusing on something.
Processing Speed	Ability to process information fast.
Working Memory	Ability to remember an exact sequence of steps in a process and the ability to orientate in this process.
Long Term Memory	Ability to learn and remember information for a long time.
ICT Literacy	Ability to use modern information technology.
Hand-Eye Coordination	Ability to coordinate the movement of the hands with things seen.
Speech Articulation	Ability to speak
Finger Precision	Ability to move the fingers precisely.
Hand Precision	Ability to move the hand precisely.
Arm Precision	Ability to move the arms precisely.
Contact Grip	Ability to control things by touching them.
Pinch Grip	Ability to press single buttons.
Clench Grip	Ability to hold object.
First Name	The first name of the user.
Last Name	The last name of the user.
Email Address	The email address of the user.
Preferred Language	The language the user prefers to use.
Successful Interactions	The number of successful interactions with the system.
State transitions	The number of state transitions the user carried out.
MyUI Experience	The experience with the MyUI system.
PreferenceTonalOutput	Selects whether the user prefers output enhanced with sounds.
PreferenceSpeechOutput	Selects whether the user prefers speech-output in addition to text.

Furthermore, not only sensors influence information within the user profile, but also the interactions the user conducts on the system. The interactions are tracked and evaluated automatically. The user is suggested to play certain games on his system in order to find out about abilities like hand precision or processing speed.

B. Device Characteristics

The device profile is created and updated by the device manager via device-specific patterns. The following Table II shows how the device profile looks like.

TABLE II. THE MYUI DEVICE PROFILE

Device variable	Description

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

category	Category of the device, on which the application is running
titleFontSize	Font size of title elements
bodyTextFontSize	Font size of all body text elements
complementaryTextFontSize	Font size of all complementary text elements, e.g., additional descriptions or comments of minor importance for the user
cellSize	The variable grid describes how many pixels are needed to build one cell of a grid.
cellCount	The variable cellCount describes how many cells are needed to fill the display.
titleBarArea	The Coordinates and layout properties of the title bar
contentArea	The Coordinates and layout properties of the content area
contentControlArea	The Coordinates and layout properties of the content control area
functionsArea	The Coordinates and layout properties of the functions area
adaptationArea	The Coordinates and layout properties of the adaptation area
microphoneAvailable	Whether there is a microphone available
touchAvailable	Whether there is a touch input device available
pointingDeviceAvailable	Whether there is a pointing device available
keyboardAvailable	Whether there is a keyboard available
remoteControlAvailable	Whether there is a remote control available
cameraAvailable	Whether there is a camera available
displayAvailable	whether there is a display available
speakerAvailable	whether there is a speaker available

Those parameters of user profile and device profile are input for the user interface generation process and will be used in order to adapt the user interface optimally to a certain user, his device and his current environment.

IV. Pattern Approach

Within the MyUI interface generation system, design patterns are used to display the knowledge about how to adapt the system to a certain user. These patterns are stored within the design patterns repository, which includes proven design solutions for optimal accessibility and usability. In order to assure flexibility and extensibility, the MyUI design pattern repository is developed on basis of a media wiki (see [12] and Figure I).

The MyUI design patterns within this repository are human-readable without any further learning or knowledge about rule-based systems. They are linked to re-usable software components. Every instance of a MyUI user interface is the result of the composition of a number of single design patterns – or clearly put: reusable software components linked to these design patterns.

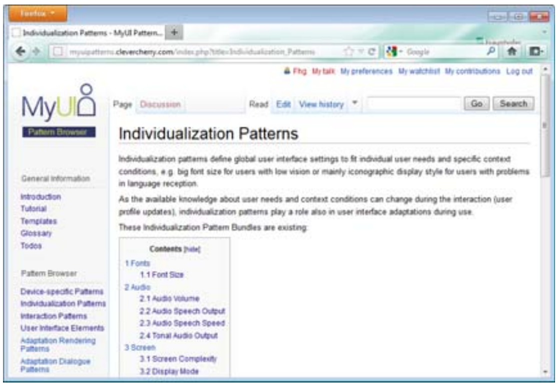


FIGURE I. The MyUI Design Pattern repository can be accessed through the MyUI Pattern Browser

The MyUI design patterns repository includes six categories of design patterns:

Device-specific patterns interprete the information provided by the device. The output is a device profile, which contains standardized information about the device (e.g. hardware provided by the device, suitable font sizes for that device, etc.)

Individualization patterns, which will be shown in detail later on in this paper, take a look on the user and his needs, as well as the context conditions. E.g. they define big font size for users with low vision or a simple, plain surface for people with attention deficits.

Interaction patterns show solutions for given interaction situations, for example, a optimized formular for an interaction situation, where a user should provide input. They use information of the user and the device profile.

User interface elements are basic elements, of which interaction patterns consist. They provide the generic primitives required composing the interaction patterns, e.g., a "selection list" interaction pattern requires "option buttons" as user interface elements.

Adaptation patterns are used, when the user interface changes, due to environmental changes or due to besser knowledge about the user, his needs and his disabilities. They define the mechanisms of switching from one instance of a user interface to another.

V. Interaction Patterns for the Adaptation of the Interface to a Certain User

Individualization patterns are the core piece of the automatic user interface individualization. They create and update the user interface profile. They are closely related to specific user characteristics as stored in the user profile and relevant device characteristics as stored in the device profile. They process the current user profile and device profile and "translate" the user, environment and device characteristics into user interface features, i.e. global settings.

Each time when changes occur to the user profile, all rules of the individualization patterns will be recalculated and the best fitting individualization patterns will be

Copyright (c) IARIA, 2012.

activated. This will result in changes to the global settings as defined in the user interface profile. Thus, user interface parameterization is a permanently on-going process during the interaction. Changes in the user profile serve as triggers for a parameterization updates.

Individualization patterns set parameters to the “User Interface Profile”. The information within this profile is used by interaction components (e.g., MainMenu) in order to decide, how to show itself (e.g., as textual menu, as graphical menu with a certain amount of entries or as a sound menu).

A. Format of Individualization Patterns

Individualization pattern bundles collect patterns that suite the same purpose. The MyUI design patterns repository includes bundle descriptions for individualization patterns in the following format:

TABLE III. Template for individualization pattern bundles

<ID>	<name of the individualization pattern bundle>
Problem	< describe a general interaction problem related to a value range(s) of certain aspects/variables of the user profile and the device profile>
sets <variable(s)> as used by <pattern bundle>	<put in, which variables are set by this bundle, connecting individualization patterns with interaction patterns by variables of the user interface profile being the output of the one and the input of the other pattern. e.g., user interface profile variables such as numericNavigation determine, which variant of the interaction pattern bundle »Main Menu« to select>
Patterns	<indicate the patterns of the bundle>

Individualization patterns are described in the following format:

TABLE IV. Template for individualization patterns

<ID>	<name of the individualization pattern>
Problem	< describe a general interaction problem related to a value range(s) of certain aspects/variables of the user profile and the device profile>
Pattern Bundle	<pattern bundle>
Context	**IF** <check values of user profile variable(s) and/or device profile variable(s)>
Solution	**THEN** Set <user interface profile variable(s)> = <value(s)>
Code reference	<if applicable, provide a reference to the related code to implement the solution>
Diagram	<if applicable, illustrate the design solution in a schematic and concise manner>
Rationale (references)	<explain the principles or rationale behind the pattern and provide references to literature, standards, etc.>
Substitutes	<individualization pattern(s)>: grouping two or more individualization patterns to a pattern bundle
uses <variable(s)> as set by <pattern bundle>	<fill in variables used within the if-Statement of this pattern, reflecting the sequential steps of device profiling and individualization and their dependencies in the UI parameterization process>
requires <variable(s)> as set by <pattern bundle>	<fill in variables used within the then-Statement of this pattern, establishing the connection to variables set by device-specific patterns, which are referred to in the solution statement of a individualization pattern>

B. Example: Individualization Pattern Bundle “Display Mode”

To show how individualization patterns are used, the pattern bundle “Display Mode” will be explained. These patterns suit the purpose to set the ratio between textual information and graphical information.

TABLE V. Pattern Bundle “Display Mode”

Individualization Pattern Bundle	Display Mode
Problem	These patterns decide about the ratio between textual information and graphical information
sets <variable(s)> as used by <pattern bundle>	**sets** *displayMode* **as used by** MainMenu, FunctionsMenu, NavigationMenu, ListWithAttributes MultiListWithAtrributes, TreeStructure
Patterns	Display Mode – mainly text (default) Display Mode – text and graphics Display Mode – mainly graphics Display Mode – graphics only Display Mode – text only

All patterns of this pattern bundle will set the parameter *displayMode*, which will be used by various Interaction Patterns like MainMenu etc.

Table VI presents an example of a pattern, which determines the ratio of textual and graphical information on the screen in accordance with current setting in the user profile variables *understandingAbstractSigns* and *languageReception*.

TABLE VI. Pattern “Display Mode – text & graphics”

Individualization Pattern	Display Mode – text and graphics
Problem	Users with mild language reception impairments will benefit from a balanced display mode with text and graphics.
Pattern Bundle	Display Mode
Context	**IF** (*understandingAbstractSigns* < 2 AND 1 ≤ *languageReception* < 2) OR (*understandingAbstractSigns* ≥ 2 AND *languageReception* ≥ 2)
Solution	**THEN** set *displayMode*: text and graphics
Rationale (references)	See MyUI Document D2.1 (p. 57), URC11 citing ER-FG-01 (2010), National Institute on Aging (NIA) National Library of Medicine (2005) Making Your Website Senior FriendlyL A Checklist. (http://www.nia.nih.gov/HealthInformation /Publications/website.htm, Kurniawan, S. and Zaphiris, P. (2005). Research-Derived Web Design Guidelines for Older People. Proceedings of 7th international ACM SIGACCESS Conference on Computers and Accessibility 20-05 (ASSETS’05), pp 129-135: http://portal.acm.org/ citation.cfm?id=1090810
Substitutes	Display Mode – mainly text Display Mode – mainly graphics Display Mode – graphics only Display Mode – text only
uses <variable(s)> as set by <pattern bundle>	-
requires <variable(s)> as set by <pattern bundle>	-

The other individualization patterns of the pattern bundle “Display Mode” work likewise.

C. *Example: Individualization Pattern Bundle Font Size*

Another interesting example for a pattern bundle is “Font Size”. This individualization patterns identify the adequate font size for a certain user, using a certain device. Beforehand, device-specific patterns would have been executed calculating possible font sizes for the device and his resolution. As it can be seen, the individualization patterns refer to these variables.

TABLE VII. Pattern Bundle “Font Size”

Individualization Pattern Bundle	Display Mode
Problem	For different situations there are different font sizes. Within this pattern bundle, the patterns set the appropriate font size according to user profile. They don't set fixed point sizes, but select relative font size variables as set by device-specific patterns
sets <variable(s)> as used by <pattern bundle>	**sets** *titleFontSize* *bodyTextFontSize* *complementaryTextFontSize* **as used by** several user interface elements
Patterns	Font Size - Philips Standard (Default) Font Size - Medium Font Size - Large Font Size - X Large

Table VIII presents an example of a pattern, which determines the font size in accordance with the current setting in the user profile variable *visualAcuity*. Beforehand, device-specific patterns have set the possible font sizes of this device, as an extra-large font size one a smartphone would probably have another value then an extra-large font size on an iTV.

TABLE VIII. Pattern “Font Size – Xlarge”

Individualization Pattern	Font Size - Xlarge
Problem	Users with significant vision impair-ments need significantly increased font sizes in order to be able to read the dis-played text in a comfortable manner.
Pattern Bundle	Font Size
Context	**IF** 3 ≤ *visualAcuity* < 4
Solution	**THEN** set titleFontSize = titleFontSizes.xlarge set bodyTextFontSize = bodyTextFontSizes.xlarge set complementaryTextFontSize = complementaryTextFontSizes.xlarge
Rationale (references)	Percentage re-sizing based on WCAG 2008. Font resizing standards for CSS suggest a scale between 100-200%
Substitutes	Font Size - Philips Standard (Default) Font Size - Medium Font Size - Large

uses <variable(s)> as set by <pattern bundle>	**-**
requires <variable(s)> as set by <pattern bundle>	**requires** titleFontSizes bodyTextFontSizes complementaryTextFontSizes **as set by** Device Font Sizes

The other individualization patterns of the pattern bundle “Font Size” work likewise.

Within the pattern repository, for every facet that influences the user interface, individualization patterns decide about how the ideal user interface for this special situation would look like. Further individualization pattern bundles, are for example:

- Audio Speech Speed: a pattern bundle that decides, depending on user characteristics like attention and processingSpeed, about the speed when vocal audio output is needed
- Navigate and Select: A pattern bundle, that considers used input devices as well as user characteristics (e.g., ICTLiteracy, fingerPrecision) in order to detect the method for navigation and selection for this certain usage situation

Detailed information about the shown pattern bundles and a lot of further patterns and pattern bundles can be found at [12].

VI. Conclusion and Future Work

The MyUI framework enables an advanced development of accessible, adaptive user interfaces through a generic approach including all relevant parameters for UI adaptation, which are user, device and environment (context). Changes are tracked automatically through system interaction and sensors, and even runtime changes on one of the parameters can be treated by a runtime change of the user interface using adaptation patterns.

The knowledge about best practices for accessible and adaptive user interfaces is stored within an open, easy readable design pattern repository [12]. Different categories of patterns address different steps within the user interface adaptation process. The individualization patterns we have shown in detail within this paper determine global characteristics of the user interface, which are constant within the whole adapted user interface.

Today, the MyUI design pattern repository includes pattern for use cases like email and instant messaging. In the future, the patterns will be refined and improved stepwise while enlarging the MyUI system to develop more applications for smart environments and devices. The repository is open for changes by the community and presents a complete overview about knowledge and intelligence of the system. As the barrier for changes on the patterns is low, new scientific findings or practical insights can easily be included and will improve the adaptation process further on.

VII. Acknowledgements

The authors acknowledge the help of other partners in the MyUI consortium in this work: Phillips Consumer Lifestyle, Ingenieria Y Soluciones Informaticas Del Sur (ISOIN), CleverCherry.com, Universidad Carlos III de Madrid, The University of Nottingham, Forschungszentrum Informatik, Birmingham City Council, Semmelweis Egyetem, and Ayuntamiento De Getafe.

The MyUI project is funded by the European Union (FP7-ICT-2009-4-248606).

References

[1] M. Blumendorf, G. Lehmann, and S. Albayrak (2010). Bridging Models and Systems at Runtime to Build Adaptive User Interfaces. InProceedings of the 2nd ACM SIGHCHI symposium on Engineering interactive computing systems (EICS ’10). ACM, New York, NY, USA, 9-18.

[2] J. Borchers. A pattern approach to interactive design. UK, John Wiley & Sons Ltd., 2001.

[3] A. Bunt, C. Conati, and J. McGrenere. (2009). A Mixed-Initiative Approach to Interface Personalization. AI Magazine 30(4).

[4] H. Dieterich, U. Malinowski, T. Kühne, and M. Schneider-Hufschmidt, (1993). State of the Art in Adaptive User Interfaces. In: M. Schneider-Hufschmidt, T. Kühne & U. Malinowski (Eds.): Adaptive User Interfaces: Principles and practice. Amsterdam: North-Holland, 13–48.

[5] G. Fischer, (1993). Shared knowledge in Cooperative Problem-Solving Systems – Integrating Adaptive and Adaptable Components. In: M. Schneider-Hufschmidt, T. Kühne & U. Malinowski (Eds.): Adaptive User Interfaces: Principles and practice. Amsterdam: North-Holland, 49–68.

[6] M. C. Huebscher and J. A. McCann (2004). Adaptive middleware for context-aware applications in smart-homes. Proceedings of the 2nd workshop on Middleware for pervasive and ad-hoc computing. ACM New York, NY, USA, 111-116.

[7] A. Kobsa, J. Koenemann, and W. Pohl (2001). Personalised hypermedia presentation techniques for improving online customer relationships. The Knowledge Engineering Review, Vol. 16:2, 111–155.

[8] W. E. Mackay (1991). Triggers and barriers to customizing software, in: S. P. Robertson, G. M. Olson & J. S. Olson (Eds.). Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’91), New York: ACM. 153–160.

[9] M. Peissner, A. Schuller, and D. Spath (2011). A Design Patterns Approach to Adaptive User Interfaces for Users with special Needs. In: HCII'11 Proceedings of the 14th international conference on Human-computer interaction: design and development approaches - Volume Part I. Springer-Verlag Berlin, Heidelberg, 268-276.

[10] F. Portet, M. Vacher, C. Golanski, C. Roux, and B. Meillon (2011). Design and evaluation of a smart home voice interface for elderly: acceptability and objection aspects, in: Personal and Ubiquitous Computing, London, vol. 15, 1-18.

[11] FZI Living Lab Ambient Assisted Living FZI Forschungszentrum Informatik, Germany, http://aal.fzi.de/, last access on 2012/03/08

[12] MyUI Design Patterns Repository. Available at http://myuipatterns.clevercherry.com, last access on 2012/03/08

Ideas for a Trust Indicator in the Internet of Things

Wolfgang Leister and Trenton Schulz
Norsk Regnesentral
Oslo, Norway
Email: wolfgang.leister@nr.no, trenton.schulz@nr.no

***Abstract*—The Internet of Things will connect many different devices. In order to realise this, users must be willing to trust the devices and communication that happens automatically. We explore the different meanings of trust and strategies that can be used to determine if something is trustworthy and propose a model for trust that takes into account people, devices, and their connections. The model uses à priori and à posteriori trust to give an indication of how much a user can trust or distrust the information provided by things. This trust indicator can inform users' decisions on whether or not to use a device or service.**

***Keywords*—*Trust*; *trust model*; *Internet of Things*.**

I. Introduction

The *Internet of Things* (IoT) refers to uniquely identifiable objects (things) and their virtual representations in an Internet-like structure. This term was first used in 1999 by Ashton [1]. Other definitions of IoT have appeared as technology progresses. A *thing* is a real or virtual object, e.g., a device or a web service, offering one or more services. Since implementations of the IoT integrate many things belonging to different actors used for various purposes, we must question to what degree we can trust these things both as a individual entities and as a federation of entities.

Even when reading the value from one isolated thing we can identify challenges regarding trust. For instance, when looking at a watch we need to consider whether we can trust the time it displays: the watch might show the wrong time for some reason. We also might be observed looking at our watch, which might breach our privacy since observers might consider us impatient or bored. This is in line with Watzlawick's first axiom that suggests that it is impossible not to communicate [2]; as a consequence, humans always reveal some information about themselves. These problems are compounded when extending trust issues to the IoT where we are dealing with many different things. It would, therefore, be helpful to have a mechanism to measure how and when to trust things on the IoT and indicate this to the user; users could then take countermeasures if needed. These measures can be based on an applicable model and theory of trust. This would allow users of IoT services to consider their actions and reduce the costs of checking trustworthiness through other frameworks.

Our work is inspired by the use of the IoT in health care where monitoring systems using various sensors and devices form a communication internetwork. These devices also need to communicate with a health care infrastructure. In health care systems, the requirements for privacy, security, integrity and availability are especially high. We have previously analysed the security model in patient monitoring systems [3] and suggested a framework for implementing sensor networks in health care [4]. We intend to extend this work with a sustainable trust model.

We posit that trust in the IoT is not transparent enough for the user. Our contribution consists of ideas for a trust model and metrics for the IoT including both channels, things, humans, and services. In order to reduce threats regarding security, privacy, and functionality when using the IoT, we distinguish between dimensions of à priori and à posteriori trust, as well as computational, technical, and behavioural trust. The developed trust model and its metrics will be used to *a*) give indications at to which security mechanisms to employ in the implementations of systems; *b*) give indications to the users at any time of how much they can trust a system, allowing users the opportunity to consider using devices in the IoT, or to use different parameters or settings; and *c*) annotate data retrieved from a system with trust values derived from the trust model.

In this paper, we first define trust and trust models in Section II before presenting our model and metrics of trust in Section III. Finally, we provide an outlook in Section IV.

II. About Trust

For the purposes of this note, we define trust as the *degree of reliance a person or thing puts in a separate thing's behaviour in a specific context*. For the IoT, trust is defined as the expectation a thing will do what it claims without bringing harm to the user [5]. This includes the perception of being secure, e.g., resilient to attacks, and that the user *a*) knows who is being spoken to, *b*) knows what is going on, *c*) feels in control of what is going on, and *d*) understands the distributed services that are involved [6].

Trust also implies that users receive information that they believe to be true and of a certain quality and timeliness. The received information can be trustworthy (usable immediately), trustworthy with alteration (usable after alteration), or untrustworthy (worthless). In the absence of trust, the user needs to consider whether it might be beneficial to abstain from using certain services of the IoT.

A. Trust

Trust is defined in several disciplines, such as sociology, psychology, ethics, economics, management, and computer science for different purposes and application areas. We distinguish between different types of trust, such as *a*) behavioural trust: expectations to the behaviour of a participant, often based on a game-theoretical approach; *b*) computational trust: the human notion of trust in the digital world, i.e., trust between agents that do not have their own agenda, such as nodes in a sensor network; and *c*) technical trust: establishing and evaluating a trust chain between devices in the IoT by means of information security technologies.

In the literature, we find a variety of definitions and categorisation of aspects of trust, such as in the review paper by Lamsal [7]. Most definitions that cover multiple disciplines define trust as the *willingness of a trustor to be vulnerable to actions of the trustee* [8] or refinements of this. See also the work by Rousseau et al. [9] and the first sections by Colquitt et al. [10]. Romano [11] defines trust as the *subjective assessment of another's influence in terms of the extent of one's perception about the quality and significance of another's impact over one's outcome in a given situation, such that one's expectation of, openness to, and inclination toward such influence provide a sense of control over the potential outcomes of the situation.*

Sabater and Sierra [12] review computational trust and reputation models. Gambetta [13] defines trust in terms of mathematics as *a particular level of the subjective probability with which an agent assesses that another agent . . . will perform a particular action. . . .* Here, trust can be quantified from *distrust*, via *no trust* to *blind trust*. Trust is only relevant if a possibility of distrust, betrayal, exit, or defection exists.

Literature using technical trust, as in the work by Fritsch et al. [6] identify the most important trust information for end users in the IoT as *a*) recognition or identification of the federation of things one connects to, *b*) ability to identify the owner, controller, or legally responsible entity behind a federation, and *c*) transparency concerning functionality, and security and privacy assurance information. We posit that this view only represents a part of the trust requirements in the IoT, and prefer the more multi-disciplinary definitions.

Cloud computing is another area where trust is a necessity. This includes security and privacy, availability, and conformance to laws from different areas (i.e., those using the cloud have strict rules about data access that the service provider must follow). Khan and Malluhi [14] detail these issues and point out that service-level agreements (SLA's) can help solve these problems. While this is true, we feel that the number of different things in the IoT will make it problematic to establish a SLA with each actor. An alternate method is necessary.

Marsh [15] formalises trust as a computational concept. He cites Deutsch [16] with the utility theory and 19 hypotheses about trust. The first hypothesis states that an individual makes a trusting choice when $Va^{+} \times S.P.^{+} > Va^{-} \times S.P.^{-} + K$, where Va^{+} and Va^{-} are positive and negative valence or utility, $S.P.^{+}$ and $S.P.^{-}$ are the corresponding subjective probabilities, and K is the security level for this individual.

Marsh presents an example heuristic formalism for trust where agents, knowledge, importance, and utility are used to define *basic trust* T_x, *general trust* $T_x(y)$, and *situational trust* $T_x(y, \alpha)$, all defined in the range $[-1, +1)$ [15, p. 59]. This notation can be used in a temporally-indexed notation. Further, Marsh [15, p. 68] outlines the cooperation threshold where the trust $T_x(y, \alpha)$ must be higher so that agent x cooperates with agent y in situation α. The cooperation threshold, in turn, is dependent on perceived risk, perceived competence, and importance; see also Coleman [17] and Gambetta [13].

B. Trust Strategies and Agent Dispositions

A trust indicator also needs to take into account the agent dispositions, such as optimism, pessimism, pragmatism, and realism [15, p. 65], in order to give suitable hints on the current trust situation. O'Hara et al. [18] lists five trust strategies in the semantic web that we can apply to things in the IoT: *1*) *optimistic strategy*: assume all agents are trustworthy unless proven otherwise, *2*) *pessimistic strategy*: assume all agents are untrustworthy unless proven otherwise, *3*) *centralised strategy*: trust information is managed by and obtained from centralised institutions, *4*) *investigative strategy*: check and evaluate agents to determine their trustworthiness, and *5*) *transitive strategy*: analyse networks of agents to determine their trustworthiness.

While all these strategies can be used in different situations in the IoT [6], the model and trust indicator that we describe here use the transitive strategy or the centralised strategy. We can trivially model the optimistic or pessimistic with any model. Our model requires other actors to properly model an investigative strategy.

C. Trust Models for the IoT

Gligor and Wing [19] present a theory of trust in networks of humans and computers that consists of elements of computational trust and behavioural trust. They present a simple communication model of entities and channels. The participants of this model can be human users, network hosts, or network applications; both honest users and machines, as well as intruders, are allowed. For human users, behavioural trust following a game-theoretical approach is used.

For computational trust, Gligor and Wing assume secure communication channels, so that they only need to consider whether a receiver (trustor) can trust the sender (trustee). In order to trust the received information, the value of information must be higher than the costs of trusting. Achieving trust can be done by verifying whether the sender can be trusted, e.g., by second opinions. However, such a second opinion might never arrive. Therefore, the receiver might be forced to use information without validation in some situations. Gligor and Wing use the concept of *isolation* that can be achieved by direct receiver verification, second opinion, etc. Trustworthiness, correctness, and the act of trusting are part of their framework.

While secure channels can be assumed for some applications, e.g., applications on the web, we cannot assume this for all networks in the IoT. Some types of sensor networks have restrictions with respect to resources, e.g., battery capacity

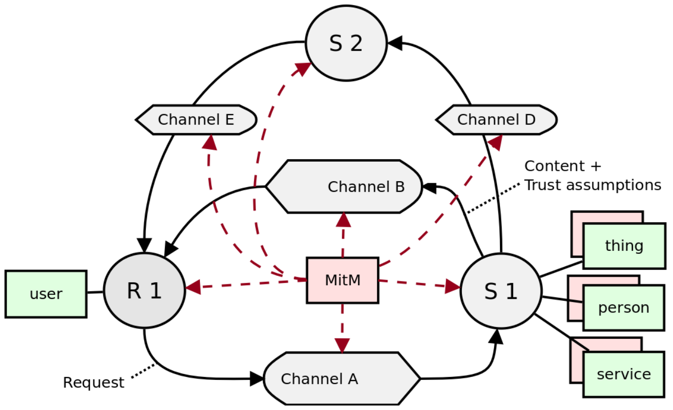


Fig. 1. Trust diagram of actors, the communication channels used, and a potential man in the middle for a scenario with one device, R_1, of the observer and two things, S_1 and S_2. S_2 offers a secondary channel from S_1 to R_1.

and power consumption, so that secure channels cannot be assumed. Therefore, we need to include the channels in our model of trust for the IoT.

III. Model and Metrics of Trust

In the IoT, we assume peers in a network that communicate with each other; peers can have roles as senders, receivers, or both. In this network, we have three types of actors: *a*) the things, devices, or services in the IoT; *b*) the persons with an intent or interest in the exchanged data; and *c*) the communication channels between things and between things and persons. Fig. 1 presents these actors and their relationships.

A. Characteristics of Actors in the IoT

Since we are interested in the specifics of the IoT, we disregard behavioural trust, and, instead, refer to the work by Gligor and Wing [19]. However, depending on who is controlling a thing, it can have its role dictated by its owner, such as being an honest node, a spy node, or otherwise compromised. Note also that things are trust-wise in a federation with the person controlling them.

To define the trust for things and channels we need to look at the possible threats that can arise from the use of the IoT. As outlined by Leister et al. [3], threats in health care applications include attacks on confidentiality, privacy, integrity, availability and non-repudiation. While these threats can be countered by technical measures, this is not always possible or practical.

For the channels, for direct communication and, indirectly, via other things, we can either *a*) use a secure channel, e.g., through cryptography and authentication, *b*) use channels with certain technical measures to counter some threats; or *c*) use unsecured channels. The threats imposed by a man-in-the-middle (MitM) attack include eavesdropping, modification of data (integrity), and lost data, e.g., through routing attacks. We can also model technical failures as a MitM attack with a random behaviour. All these are side-effects of using the IoT that users do not expect and thus breach trust.

Things need to have an identity in order to be assigned a trust value. Also services in the IoT need to be identifiable for this purpose. Consider services as a federation of things where the owner of a service needs to be identifiable; attacks on things can cause the same threats as for channels. Additionally, there might be deviations in data due to setup errors or for other technical reasons. For example, a clock that can experience a drift when not set regularly or a thermometer that is accurate but is affected by unwanted environmental influences. Another cause could be known biased actors, such as certain weather forecasts.

When presenting a too low trust value for a thing or channel to a user, the user can *a*) not send information, *b*) not request a value, *c*) employ technical measures to secure the value, *d*) discard a received value, *e*) delay usage or store a received value (and attached information) without using it until more evidence is collected, or *f*) adjust a received value by a suitable offset, e.g., from past experience.

B. A Trust Model for the IoT

As defined by Marsh [15], we use trust in the range $[-1, +1]$, the value 1 indicating complete (blind) trust[1], the value -1 indicating complete distrust, and the value 0 indicating indecision (i.e., more information is needed).

Let the observer be R_1. For the observer R_1, we define the *à priori trust* $-1 \leq \tau^{\triangleright}_{S_1} \leq 1$ as the trust before sending a message to an object; this trust consists of the trust of the Channel A and of the object (thing) T_1:

$$\tau^{\triangleright}_{S_1} = \tau^{\triangleright}_{A} \cdot \tau^{\triangleright}_{T_1}$$

The à priori trust can be used to decide whether to send the message, or not, with the help of a threshold value. Note that in most cases the trust value of both the onwards and the return path need to be considered since most requests will result in a sequence of responses from a node to the receiver.

For the observer R_1 we define the *à posteriori trust* $-1 \leq \tau^{\triangleleft}_{S_1} \leq 1$ as the trust after receiving a message from an object; this trust consists of the trust of the Channel B and of the object (thing) T_1:

$$\tau^{\triangleleft}_{S_1} = \tau^{\triangleleft}_{B} \cdot \tau^{\triangleleft}_{T_1}$$

The à posteriori trust is the trust value to decide whether a message can be used. The à posteriori trust comes with an adjustment function $a^{\triangleleft}_{S_1}$ that can be applied to the received message. The adjustment function contains components from Channel B and the object T_1. The à posteriori trust is the trust value *after* having applied the adjustment function to the value $v^{\triangleleft}_{S_1}$. In order to make a decision what to do with a value, R_1 receives and stores the tuple $(\tau^{\triangleleft}_{S_1}, a^{\triangleleft}_{S_1}, v^{\triangleleft}_{S_1})$.

When using these metrics to indicate trust, we need to look into *a*) the factor of time that temporally indexed notations need to be applied and *b*) the transitive behaviour. For the transitive behaviour, we need to study how a chain of nodes and channels behave as long as $\tau > 0$. Note that in the case of distrust, i.e., $\tau < 0$, we cannot set up a useful assumption

[1]The term *blind trust* poses philosophical problems; therefore, most authors set the trust range to $(-1, +1)$. For practical reasons, initially, we allow a trust value of exactly 1.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

for the trust of the entire chain. Using technical measures to counter security and privacy threats we can obtain $\tau < 0$ while still the threat exists that a result will not arrive at the receiver due to some attack on the availability. Therefore, the trust estimates should be split up into separate values for security, privacy, availability, and so on.

IV. Outlook

We presented our ideas towards a formalism that tries to describe trust in the IoT. Issues regarding chains of nodes and channels and composition of nodes and channels to complex networks are challenges to such a framework. Another challenge could be perceived trust as a relation to the Quality of Experience (QoE) a user experiences from a service (e.g., distrusting a thing because the result arrived late or that the result is sometimes inaccurate). We envisage that trust values from measurements and user assessments can be established in a similar way as described by Leister et al. [20] for measured and perceived video quality (i.e., QoS vs. QoE).

On the practical side, another challenge to investigate are ways the trust indicator can best be presented to users in an understandable way. These could include a traffic light metaphor or a more elaborate dashboard where trust values and suggestions to the user are presented. There must be multiple ways to present this as some things may not have a display. We are confident that a formalism will help to create indicators that can guide users of services and data in the IoT, and that they can experience the benefits of the IoT rather than suffering from bogus services and unwanted information leaks.

Acknowledgements

This research is funded in part by the EU project FP7-258360 uTRUSTit: *Usable Trust in the Internet of Things*.

References

[1] K. Ashton, "That 'Internet of Things' Thing," *RFID Journal*, 22 July 2009, accessed March 9, 2012. [Online]. Available: http://www.rfidjournal.com/article/view/4986

[2] P. Watzlawick, J. H. Beavin, and M. D. D. Jackson, *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes*. W. W. Norton & Company, Mar. 1967.

[3] W. Leister, T. Fretland, and I. Balasingham, "Security and authentication architecture using MPEG-21 for wireless patient monitoring systems," *International Journal on Advances in Security*, vol. 2, no. 1, pp. 16–29, 2009.

[4] W. Leister, T. Schulz, A. Lie, K. Grythe, and I. Balasingham, "Quality of service, adaptation, and security provisioning in wireless patient monitoring systems," in *Biomedical Engineering, Trends in Electronics, Communications and Software*, A. N. Laskovski, Ed. Intech, 2011, ch. 36, pp. 711–736.

[5] I. N. Zoltán Hornak, D. Petró, J. Schrammel, P. Wolkerstorfer, L. Ellensohn, A. Geven, K. Kristjansdottir, L. Fritsch, T. Schulz, H. Abie, F. Pürzel, and V. Wittstock, "D.3.1 uTRUSTit Technology and Standard Report," uTRUSTit - Usable Trust in the Internet of Things, EU project deliverable D3.1, 2010.

[6] L. Fritsch, A.-K. Groven, and T. Schulz, "On the Internet of Things, Trust is Relative," in *Privacy, Trust and Interaction in the Internet of Things*, ser. Lectures in Computer Science, vol. 7040. Springer, 2011, pp. 263–269, in press.

[7] P. Lamsal, "Understanding trust and security," Online-Reference, October 2001, accessed March 9, 2012. [Online]. Available: http://www.cs.helsinki.fi/u/lamsal/asgn/trust/UnderstandingTrustAndSecurity.pdf

[8] R. C. Mayer, J. H. Davis, and F. Schoorman, "An Integrative Model of Organizational Trust," *Academy of Management Review*, vol. 20(3), pp. 709–734, 1995.

[9] D. M. Rousseau, S. B. Sitkin, R. S. Burt, and C. Camerer, "Not so different after all: A cross-discipline view of trust," *Academy of Management Review*, vol. 23, no. 3, pp. 393–404, 1998.

[10] J. A. Colquitt, B. A. Scott, and J. A. LePine, "Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance." *J Appl Psychol*, vol. 92, no. 4, pp. 909–27, 2007.

[11] D. M. Romano, "The nature of trust: Conceptual and operational clarification," Ph.D. dissertation, Louisiana State University and Agricultural and Mechanical College, 2003.

[12] J. Sabater and C. Sierra, "Review on computational trust and reputation models," *Artif. Intell. Rev.*, vol. 24, pp. 33–60, September 2005.

[13] D. Gambetta, "Can We Trust Trust?" in *Trust: Making and Breaking Cooperative Relations*, D. Gambetta, Ed. University of Oxford, 2000, ch. 13, pp. 213–237.

[14] K. M. Khan and Q. M. Malluhi, "Establishing trust in cloud computing," *IT Professional*, vol. 12, no. 5, pp. 20–27, 2010.

[15] S. P. Marsh, "Formalising trust as a computational concept," Ph.D. dissertation, University of Stirling, April 1994.

[16] M. Deutsch, *The Resolution of Conflict*. New Haven and London: Yale University Press, 1973.

[17] J. S. Coleman, *The Foundations of Social Theory*. The Belknap Press of the University of Harvard, 1990.

[18] K. O'Hara, H. Alani, Y. Kalfoglou, and N. Shadbolt, "Trust strategies for the semantic web," in *ISWC 3rd International Workshop on Trust, Security, and Reputation on the Semantic Web*, 2004.

[19] V. Gligor and J. M. Wing, "Towards a theory of trust in networks of humans and computers," in *Proc. 19th International Workshop on Security Protocols*, ser. LNCS. Springer Verlag, March 28-30 2011.

[20] W. Leister, S. Boudko, and T. H. Røssvoll, "Adaptive video streaming through estimation of subjective video quality," *The International Journal on Advances in Systems and Measurements*, vol. 4, no. 1, pp. 109–121, 2011.

RESTful Platform Broker: A Novel Framework for Engineering Programmable Smart Spaces

Govind Raj Pothengil
Centre For Development of Advanced Computing
Noida, UP, India
pgovindraj@cdac.in

Subrat Kar, Hari Mohan Gupta
Department of Electrical Engineering,
Indian Institute of Technology,
New Delhi, India
subrat@ee.iitd.ac.in, hmgupta@ee.iitd.ac.in

***Abstract*—Designing and developing a Smart Space is a difficult engineering and research problem. This paper presents the design and implementation of a novel framework for engineering Programmable Smart Spaces called RESTful Platform Broker. The RESTful Platform Broker can be considered as a middleware on top of a platform (e.g., Android). It views the network of the Smart Space as a system bus and exposes resources (e.g., Mobile Phone cameras, Smart Phone sensors, etc.) that are present in the platform as RESTful resources. Once the resources are exposed, these can be accessed by other devices present in the Smart Space. This facilitates programming the Smart Space and developing new services from the existing ones. Further, RESTful Platform Broker uses HTTP as a glue to interconnect heterogeneous components within the Smart Space. This enables any programming language, or any tool, which can understand HTTP to manipulate resources and program the Smart Space. This is an advantage as compared to earlier approaches, which warranted the use of specific languages to program a Smart Space or use protocols like 9P/9P2000, which are not as popular as HTTP. Further, the RESTful Platform Broker uses Hypermedia as an engine for application state, which allows various components within the Smart Space to evolve independently with respect to each other, thereby, increasing Smart Space extendability. Therefore, the RESTful Platform broker eases the engineering of a Smart Space by hiding the complexity of the physical infrastructure by providing a standard interface which can be used by Smart Space programmers to extend and program it.**

Keywords-Smart Space; Pervasive Computing; REST.

I. Introduction

Pervasive computing offers some unique research and engineering challenges [1]. One of the identified research areas in Pervasive computing is a Smart Space. A Smart Space is a physical space consisting of heterogeneous computing devices and services together with a software infrastructure. The users can interact with these devices in an unobtrusive manner through the infrastructure. This software infrastructure abstracts the boundaries of the individual devices, present in the physical space. It allows the user of a Smart Space to interact with the set of devices as a whole, rather than individual devices. In other words, services, applications and peripherals in a particular device can be accessed or used by other devices easily. Most of the first generation Smart Spaces were custom built to implement a particular scenario (e.g., Smart House, Smart Hospital, Smart Meeting Room, etc.). It was difficult to incorporate new technologies and extend the Smart Space by adding new devices and services. This was mainly because these Smart Spaces lacked the ability of being programmed. The Gator Tech Smart House [2] was among the first Smart Spaces that introduced the concept of Programmable Pervasive Spaces. Programmable pervasive spaces are Smart Spaces which exist both as a runtime environment and a software library. Smart Space Programmers can, therefore, extend and evolve the Smart Space by incorporating new devices and services.

However, engineering Smart Spaces is a challenging and difficult task. The basic research challenge in engineering Smart Spaces is in the system level architecture and component level synthesis [1]. This research work aims at designing and developing a framework using which the engineering of a programmable Smart Space can be simplified.

Most of the tools that are available to engineer Smart Space provide a middleware. These middlewares provide APIs so that programmers can use it to incorporate new services and devices into the Smart Space. Though Middleware APIs provide a way to extend the Smart Space, the requirement that all devices must have API compatibility, could be a serious hindrance in extending it. Further, many projects like GAIA [4] also mandated the use of specific programming languages to extend the functionality of a Smart Space. Plan B [11] provided a different approach. It was designed as an OS which exported resources within a device as a filesystem. Any language can therefore be used to program the Smart Space. Though, this alleviated the problems with APIs, however the limitation of Plan B was that it could not be ported easily to different platforms as it was an OS. In order to overcome these limitations, a new framework for engineering programmable Smart Spaces called the RESTful Platform Broker has been designed and developed. It is designed as a middleware, so that it is easier to port to different platforms. Further, it provides a Service Oriented filesystem based mechanism to program a Smart Space so that it is not limited to a particular set of API's or specific languages.

This paper presents the design and implementation of

Copyright (c) IARIA, 2012.

RESTful Platform Broker. It helps engineers and researchers to develop programmable Smart Spaces. The main objective of the framework is to hide the complexities of the underlying infrastructural elements of a Smart Space and provide a uniform view of the resources present in it in such a way that it is easy to program and extend the Smart Space. The RESTful Platform Broker is considered as a framework because Smart Space applications can be built using it. From an implementation point of view, the RESTful platform broker is a middleware that provides an abstraction between a Smart Space programmer and the individual infrastructure elements present in the Smart Space. It views the network of the Smart Space as a system bus and exports resources present within a device/platform onto the Smart Space network as service oriented file systems. The concept of service oriented file system was proposed by Eric Van Hensberg [3]. Essentially, the RESTful Platform Broker queries the platform for resources. These resources can be in the form of hardware, e.g., Camera, or a software, such as a Text to Speech utility. Once a resource is found, it wraps the resource as a service oriented file system and exposes it as a RESTful resource, which can be accessed using HTTP. The user, however, can specify the resources they want to expose to the Smart Space and thus, control the visibility of the resources within it. The exposed resources can then be used by other devices running on other platforms within the Smart Space by issuing HTTP commands like GET, POST, DELETE and PUT. Using this framework, programmers can extend and program service compositions using any language. Further, any device that can understand HTTP can access the resources exposed by RESTful Platform Broker.

The RESTful Platform Broker has been implemented over the Android Platform and has been installed on commercially available Samsung Galaxy Pop GT-S5570 mobile phone. Interaction of RESTful Platform Broker with GNU/Linux (Ubuntu 11.04) has been achieved. The experimental setup for this has been discussed in this paper. Though the implementation of the RESTful Platform Broker is done on the Android Platform, it can run on any platform on which there is a suitable JVM. The design of the RESTful Platform Broker is such that porting to a different platform can be done with minimal changes to its codebase.

This paper is divided into 6 Sections. Section 1 introduces and presents a background of the research work. Section 2 presents the work which is related to the current research. Section 3 presents the Design and implementation of the RESTful Platform Broker. This section also presents a DNS based Service Discovery mechanism implemented as a part of this project. Section 4 presents usage scenarios of the RESTful Platform Broker. Section 5 presents results on RESTful Platform Brokers performance. Section 6 concludes the paper.

II. Related Work

The main aim for our research work is to design and develop a pervasive computing framework which will help engineers and researchers to develop programmable Smart Spaces. This section focuses on projects, using which one can design and develop pervasive computing application. This section also compares their approach with that of ours.

Gaia [4] is a Meta operating system for Smart rooms. The goal of this project was to design and develop a middleware operating system (Meta operating system), that manages the resources in a Smart Space. Conceptually it is similar to a traditional operating system which manages tasks that are common to all applications programs. Gaia extends typical operating system concepts to include context awareness, provide location awareness, detect when new devices are spontaneously added to the Smart Space and adapt when data formats are not compatible with output devices. From an architecture point of view, Gaia was implemented as a CORBA middleware. System components were implemented as distributed objects with CORBA IDL interfaces. To some extent, Gaia supports Smart Space programmability. Using Gaia, programmers can integrate devices and create new services in a Smart Space. However, Gaia [4] requires special programming language like Olympus [5] to program the Smart Space and utilize specialized mechanisms like Microservers [6] to induct a device onto a Smart Space. In contrast, in our approach any language that can handle HTTP can be used to program the Smart Space. Further, any device that supports HTTP can become a part of the Smart Space without implementing any specialized software like Microservers.

Stanford Universitys initiative, iROS [7][8] is an open source software platform for designing and developing Smart Spaces. Like Gaia, iROS can also be considered as a meta operating system for Smart Spaces. The programming model of iROS is to ensemble independent entities in the Smart Space through the use of its subsystems. iROS consists of three subsystems: EventHeap for application coordination, DataHeap for data movement and transformation and ICrafter for user control of resources. A Smart Space designed and developed using iROS as the platform needs to wrap its resources using ICrafter to function within a Smart Space. When using RETSful Platform Broker, devices need to interact through a more generic HTTP based Domain Application Protocol rather than specific mechanisms like ICrafter. This enables integration of more devices into the Smart Space, as greater numbers of devices are using the HTTP protocol [9].

EQUATOR Component Toolkit (ECT) [10] provides a mechanism for constructing pervasive computing applications by integrating sensing devices (inputs from sensors, such as *phidgets*, *motes* and *d.tools* boards) and actuation devices (of physical actuators including X10 modules, output of Internet Applications, etc.). The project uses a visual graph based editor to allow run-time interconnection of modules.

Plan B OS can be considered as a software infrastructure for Smart Spaces [11]. Plan B is an operating system based on Plan 9 [12][13] which exports the resources (both hardware and software of a device) as synthetic file systems. These resources can then be shared with other devices in the Smart Space using 9P protocol and can be used through file system operation. Our research prototype shares a common design principle with Plan B, which is to expose resources as file systems. However, our approach relies on HTTP, which is a very common protocol whereas Plan B relies on 9P/9P2000 for sharing resources. Implementing 9P protocol even on open source operating system like GNU/Linux is challenging [14]. Moreover PlanB has been implemented on desktop systems, whereas, the RESTful Platform Broker has been implemented on top of Android, which is a mobile platform. Such a platform is more apt for developing Smart Spaces as mobility of devices is one of the key characteristics of a Smart Space. Further, PlanB is an operating system and hence it is difficult to port as compared to RESTful Platform Broker which is implemented as a middleware.

Nokia Research has designed and developed a RESTful framework [15] for Smart Spaces. It supports resource discovery, authorizing access to resources with group-based security and sharing context information on a device with other devices in the Smart Space. Though their research also uses REST paradigm, they have not reported the use of Hypermedia in their architecture. RESTful Platform Broker uses Hypermedia as an engine for application state and hence has a higher level of maturity in accordance with the 3 Level REST Maturity Model (RMM), which was proposed by Richardson [16]. The advantage of using Hypermedia as an engine for application State is that, all the operations supported by resources need not be presented to a client up front. They get to know them, as they interact with the resources and are presented to them during the appropriate time. This allows better separations between the client and server component within a Smart Space and allows them to evolve independently of each other.

It is also remarked that Google has recently designed a framework called Argos [17], which aims at building a Web Centric Application Platform on top of Android. Argos allows developers to access the components of the platform (e.g., Media Player, Camera, Sensors, etc.) by using Java Script instead of Java Programming Language. It does not, however, expose Android resources over the network as done by the RESTful Platform Broker. Further, the aim of the project is to tap a large potential of developers who are trained in Web Application programming and are familiar with scripting languages such as JavaScript but not Java.

III. RESTful Platform Broker

The RESTful Platform Broker is designed and developed using the REST paradigm [18]. REST is an acronym for Representational State Transfer. REST by itself is not an architecture, but a set of constraints, which, when applied, leads to a System Architecture. In order to design or architect a RESTful system, a system designer/architect should visualize the system through the following abstractions:

Resource: A resource is a fundamental building block of a RESTful architecture. A resource is anything that is exposed by the system and is addressable. By addressable, we mean that it can be accessed by other components of the system. Using the REST paradigm, anything can be modeled as a resource and can be manipulated by the components within the system. For example, a camera in a mobile phone, a media player, the reading of a sensor, etc., can be considered as a Resource.

Representation: All the identified resources in the system will have a representation. The representation of a resource is the data being transferred from a server to a client and vice-versa. The representation contains the state of the Resource at a given point in time.

Uniform Resource Identifier (URI): A Uniform Resource Identifier can be considered as a Hyperlink to a resource. By accessing the URIs, the clients and servers can exchange Representations of Resources in the system.

The major advantage of REST architecture as compared to a typical RPC based architecture is that by using REST principles, one can organize a very complex application into simple resources. Once the resources are identified, all operations on these resources can be broken down into Create, Read, Update and Delete (CRUD) operations. This simplicity makes it easy for new clients to use an application, even if it was not specifically designed for them.

The RESTful Platform Broker is implemented as a middleware on top of the Android Platform. It views the network of the Smart Space as a system bus and exposes resources present on the Android platform like Camera, Media Player, Sensors, etc., onto a Smart Space network as a Service Oriented Filesystem [3]. Service Oriented Filesystem provides a Service Oriented Interface (usually based on REST) to the resources in a system. These resources can thus, be accessed by other devices present in the Smart Space. This systemic view has been shown in Figure 1.

Each of the exported resource has a URI associated with it. For example, the Accelerometer Sensor present on the Android Platform may have a URI as shown in [24]. Once the resource has been exposed onto the network, it can be used by other devices, which are present in the Smart Space through its URI . The interaction between a device and a resource exposed by the RESTful Platform Broker is through HTTP. For example, sending a HTTP GET to URI of the Accelerometer will return the current reading of the

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

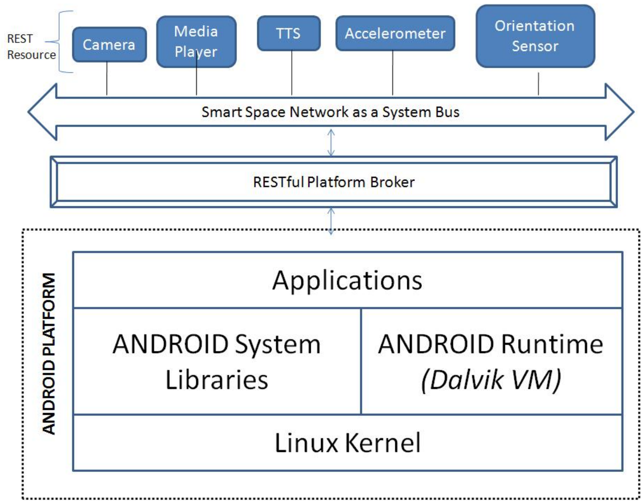


Figure 1. Systemic view of RESTful Platform Broker

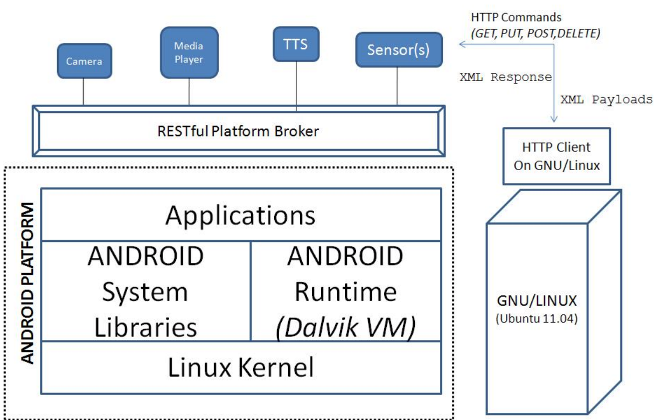


Figure 2. Interaction of RESTful Platform Broker with GNU/Linux

accelerometer. This way, resources are not confined to the physical boundaries of a device, but are available throughout the Smart Space.

Therefore, through the use of the RESTful Platform Broker, devices which may not have a particular resource may use a resource which is present in some other device. The interaction of RESTful Platform Broker with a GNU/Linux system is show in Figure 2.

RESTful Platform Broker uses Hypermedia as the engine of application state (*HATEOAS*). A Hypermedia system is characterized by the transfer of links in the resource representations, which are exchanged between the client and the server. These links advertise other resources that are participating in the application protocol. For example, consider a media player resource which is exposed by the RESTful Platform Broker. An initial HTTP GET on its entry point, e.g., [26] would result in an XML File as shown in Listing 1.

The XML in Listing 1 is an XML representation of the media player resource. The representation has a semantic markup definition provided through *link* and *rel* tags. Consider the first *link* tag, which points to the URI of Media Player Resource itself. The associated *rel* tag has the value `self`. This represents that the URI [26] can be accessed via a HTTP GET to retrieve the latest representation of the Media Player resource. It is remarked that in RESTful Platform Broker, a resource representation will always have a link to itself. This is because when using GET, we always get the latest state of a resource. For example, a subsequent GET to the URI [26] may result in an XML with the status value as Playing instead of Idle. This information can be used by layers above the platform broker to make decisions. For example, a client program can view the status of the media Resource and make a request to playa file only if it is idle. The XML representation of the MediaPlayer Resource has link to another resource namely the SongList.

The SongList Resource can be accessed using HTTP GET, which is inferred by looking at the value of the *rel* tag. A GET to the URI [25] results in an XML representation of a Song List. The XML representation provides the information of the current state and also provides a mechanism to advance to the next state. By reading the XML representation mentioned in Listing 1, a consumer can infer that there are two states possible from the current state namely; getting the representation of the media player by issuing a HTTP GET to the URI [26] or, navigate to get the Song List by issuing a HTTP GET on the URI [25]. A consumer of a resource can therefore, advance through the Domain Application Protocol of the resource exposed by the RESTful Platform Broker. When a consumer issues a HTTP GET on the URI [25], the RESTful Platform Broker generates the following XML Representation of the Song List as shown in Listing 2.

Listing 1. Response of RESTful Platform Broker on the initial GET of the Media Player Resource

```
<mediaresource
xmlns=http://platformbroker.iitd.ac.in
xmlns:mediadap="http://schemas.platformbroker.iitd.ac.in/mediadap">

  <mediadap:linkmediaType="application/vnd.platformbroker+xml"
  uri=http://ipAddress/mediaPlayer/ rel="self"/>

  <mediadap:linkmediaType="application/vnd.platformbroker+xml
  uri="http://ipAddress/mediaPlayer/songList/"rel="self"/>

  <status>Idle</status>

</mediaresource>
```

The XML representation shown in Listing 2 of the song List shows different URIs of songs and their associated *rel* values which provide the semantic markup definition.

Listing 2. XML Representation of Song List

```
<songList xmlns="http://platformbroker.iitd.ac.in"
```

```
xmlns:mediadap=”http://schemas.platformbroker.iitd.ac.in/mediadap”>

<mediadap:linkmediaType=”application/vnd.platformbroker+xml”
uri=http://ipAddress/mediaPlayer/songList/ rel=”self”/>

<song mediadap:linkmediaType=”application/vnd.platformbroker+xml”
      uri=”http://ipAddress/mediaPlayer/songName1/ rel=”self”/>

<song mediadap:linkmediaType=”application/vnd.platformbroker+xml”
      uri=”http://ipAddress/mediaPlayer/songName2/ rel=”self”/>

</songList>
```

The XML presented in Listing 3 is generated when a consumer issues a GET to URI [27]. In this XML, apart from *self*, a song's *rel* tag points to `service.put`. By accessing this URI, a consumer can play the Song mentioned in the URI by using HTTP PUT along with a suitable payload, as defined in the platform broker media Type. Therefore, through the use of Hypermedia, a client can proceed ahead to relevant states of a resource by following the corresponding URIs. This is very similar to the way in which a user navigates the World Wide Web. The advantage offered by Hypermedia is therefore loose coupling between the client and server components of the Smart Space. This allows the server components to change its rules and expand available states as needed without affecting the clients.

Listing 3. XML Representation of Song

```
<songxmlns=”http://platformbroker.iitd.ac.in”
xmlns:mediadap=”http://schemas.platformbroker.iitd.ac.in/mediadap”>

<mediadap:linkmediaType=”application/vnd.platformbroker+xml”
  uri=”http://ipAddress/mediaPlayer/songName1/”
  rel=”self”/>

<mediadap:linkmediaType=”application/vnd.platformbroker+xml”
  uri=”http://ipAddress/mediaPlayer/play/songName1”
  rel=”service.put”/>

</song>
```

A. Implementation

The implementation of the RESTFUL Platform Broker is given in Figure 3. The implementation is done on top of the Android Platform. A small footprint HTTP Server has been placed on top of Android. The HTTP server has been implemented as a pure Java based HTTP Server and can be treated as a Virtual Appliance on top of the Dalvik Virtual Machine. The RESTful Platform Broker has a layered architecture and is implemented as a hosted service on top of this HTTP server. The system's core layer consists of the Namespace Generator, which generates a namespace entry of a resource by the information given in the Namespace.xml. The Namespace.xml indicates the URIs of the Resources. The advantage of separation of URIs from

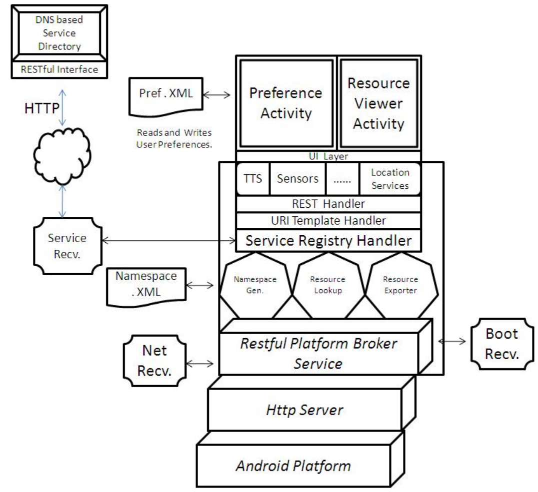


Figure 3. Implementation of RESTful Platform Broker

the RESTful Platform Brokers codebase is that the URIs of the resource can be changed without changing the code. Another component at this layer is the Resource Lookup Manager. The Resource Lookup Manager is a component, which queries the Platform for resources. Once a resource is found it works with the Namespace Generator and the Resource Exporter to export the Resource. The components of this layer of the RESTful Platform Broker, namely the Namespace Generator, Resource Lookup Manager and Resource Exporter are implementations of specific interface *INamespaceGenerator*, *ResourceLookUpManager* and *IResourceExporter* respectively. Throughout the implementation of Platform Broker, programming to an interface paradigm has been followed. This paradigm is useful as it allows multiple implementation of the same concept. For example, the *IResourceLookUpManager* can have different implementation depending on the Platform, it is being implemented for. This helps in easier porting of the RESTful Platform Broker to other Platforms.

The REST Handler and the URI Handler are handlers for the HTTP commands namely GET, PUT, POST and DELETE. These HTTP commands, when sent to the resources (e.g., TTS, Sensors, etc.), are handled by the REST Handlers and URI Handlers for these resources. The difference between the REST handler and the URI handler is as follows. The REST handler can be considered as a singleton class [19] which provides the default behavior to the HTTP Commands whereas, each URI handler for a resource can be multiple for a resource, each implementing a specialized functionality of the resource. For example, the Media Player resource may support two different URIs; one for managing the media file and another for enumerating the media files within the system.

The Service Registry Handler is the component that handles Service Registry in the Platform Broker. It interacts

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

with DNS based Service Registry to register and query for services that are registered with it. It also interacts with the Service Receiver. The Service Receiver is a broadcast receiver, which informs the Service Registry Handler of new services that have been registered. The RESTful Platform Broker has got two User Interface (UI) components. First, the Preference Activity UI, through which the user can specify the list of resources that can be exported onto the network. Second, the Resource Viewer UI, which shows a list of resources that have been found by Resource Lookup Manager and are available throughout the Smart Space. The Preference Activity writes the list of resources that the user wants to share across a Smart Space persistently in a file called Prefs.xml. The Platform Broker reads this file before exporting the resources to the Network. Further, the user can change the preference of exporting a resource at any point of time. Any such change triggers the export mechanism to reflect the changes in the preferences. This change also triggers the Service Registry Handlers to make suitable changes in the DNS based Resource Registry. Mandatory Access Control (MAC) mechanisms are implemented within the RESTful Platform Broker to prevent the use of resources which are not exposed by the user. The RESTful Platform Broker service interacts with two kinds of Broadcast Receivers namely the Network Receiver and the Boot Receiver. The Network Receiver is an important component especially from the mobility point of view. The Network receiver informs the Platform Broker, when the network is not available. This allows the Platform Broker to stop its activities, thereby conserving the battery of the Mobile device, on which the Platform Broker is installed. Further, when the Network strength is getting weak, it can de-register resources from the DNS Resource Registry. Similarly, when the device moves from a location where there is no network, to a place where network is available, the Network Broadcast Receiver informs the Platform Broker, to start its activities. The Boot receiver is used to start the RESTful Platform Broker, whenever the Android Platform is booted. These are examples of context aware adaptation that has been implemented in the RESTful Platform Broker.

The DNS Server is not a part of the RESTful Platform Broker, but has been implemented separately to provide a centralized mechanism for service registry and discovery. The implemented DNS Server provides a RESTful interface for service registration and service discovery. The RESTful Platform Broker interacts with the DNS Server by using HTTP constructs. The RESTful Platform Broker uses HTTP POST to register a service with DNS Server. The HTTP POST request to register a resource, consists of an XML payload that specifies the resource to be registered. The XML payload consists of the details of the SRV, PTR and TXT records that are required by the DNS Resource Registry. The details of the records are mentioned in Section III-B. The RESTful Platform Broker uses HTTP GET to get information of the available resources within the Smart Space, HTTP PUT to update a resource and HTTP DELETE to remove a resource from the resource registry.

Listing 4. SRV Record of DNS

```
_pb_http._tcp.local IN SRV 0 0 80 GovindPhone.local
```

Listing 5. DNS Reply to Query

```
_pb_http._tcp.local IN SRV 0 0 80 GovindPhone.local
_pb_http._tcp.local IN SRV 0 0 80 GovindLaptop.local
_pb_http._tcp.local IN SRV 0 0 80 GovindTablet.local
```

B. DNS Based Service Discovery

The RESTful Platform Broker supports DNS based service discovery. DNS Service Discovery is a way of using standard DNS programming interfaces, servers, and packet formats to browse the network for services [20]. Since, Hypermedia has been introduced as an engine for application state within the RESTful Platform Broker; a client only need to know the entry URI of a resource. Other URIs are presented as the user uses the resource. For example, a client may only need to know the URI of a media player which may be available in a particular device. Once a client issues a GET command onto this URI, other related URIs relating to usage of the Media player like getting a Song List, Playing a song, Pausing a song, Stopping a song, etc., is presented as and when needed.

A DNS Server can be used to associate various kinds of Resource Records on a particular domain. Apart from records like A-record (for address lookup) or MX-records (for mail server records), DNS also defines resource record types SRV (used to provide location and port for service instances.), TXT (Text Record, used to provide additional meta data about service instances) and PTR (Pointer Record, used to map service types to named service instances). A DNS can be used for service discovery by a combination of SRV, TXT and PTR records. SRV, TXT and PTR records are described below.

Listing 6. DNS SD Configuration

```
_camera_pb_http._tcp PTR GovindPhoneCamera._pb_http._tcp.local
_camera_pb_http._tcp PTR GovindTabletCamera._pb_http._tcp.local
_camera_pb_http._tcp PTR GovindLapCamera._pb_http._tcp.local
_mPlayer_pb_http._tcp PTR GovindPhoneMPlayer._pb_http_.tcp.local
_mPlayer_pb_http._tcp PTR GovindLapMPlayer._pb_http._tcp.local
_tts_pb_http._tcp PTR GovindPhoneTTS._pb_http._tcp.local
_accmeter_pb_http.tcp PTR GovindPhoneAccmeter._pb_http._tcp.local

GovindPhoneCamera._pb_http._tcp SRV 0 0 80 GovindPhone.local.
                                TXTpath=/camera
GovindTabletCamera._pb_http._tcp SRV 0 0 80 GovindTablet.local.
                                TXTpath=/camera
GovindLapCamera._pb_http._tcp SRV 0 0 80 GovindLap.local.
                                TXTpath=/camera
GovindPhoneMPlayer._pb_http SRV 0 0 80 GovindPhone.local.
                                TXTpath=/mediaPlayer
```

SRV Resource Records: SRV Resource Records are used to provide information about host and port within a zone on which a service is available. As an example, the SRV record for a camera resource is given at Listing 4. The line in the Listing specifies that a Platform Broker is accessible at port 80 on the device whose name is GovindPhone.local. The *.local* is a pseudo top level domain identifying a local domain. The Listing 5 shows a reply to a DNS query. This shows that Platform Broker is running on three devices namely GovindPhone, GovindLaptop and GovindTablet at port 80.

PTR Records: The SRV records have a limitation that they cannot be used to configure named instances of a service type. Further, they only support a single service for any given host and port combination. For example, just by using SRV records, we may not be able to specify all the resources that Platform Broker has exported on a specific device. Therefore, PTR records are used to map service type names to service instance names. The configuration snippet displayed in Listing 6 shows that, on the left hand side of each PTR line, a service type domain name is given and on the right hand side, a corresponding instance of that type is given.

TXT Records: The TXT Records are in the form of a key value pair. In order to use a resource exposed by the RESTful Platform Broker, an entry point of the resource is required. Once this entry point is discovered, other related URIs are presented by the RESTful Platform Broker as and when needed. TXT Records is used for providing an entry point into the resource. The example configuration in Listing 6 specifies that the service instance GovindPhoneCamera._pb_http._tcp can be accessed on GovindPhone.local. at port 80. The TXT record specifies path parameter, which the client must use for constructing the entry URI of the service. In this case, the resource GovindPhoneCamera._pb_http._tcp can be accessed by the URI [28].

The configuration file snippet below in Listing 6 provides PTR, SRV and TXT records for the various resources registered with the DNS Server of the Smart Space.

IV. Application Development Using RESTful Platform Broker

Wireless Accelerometer Based Mouse: This scenario shows the usage of reading sensor information which is available in Android on a Linux Platform. The sensor in this case is an accelerometer. The sensor reading is then used to move a mouse on the laptop by tilting the Android Phone. To construct an accelerometer based mouse using the RESTful Platform Broker, an HTTP client is designed and developed which issues a GET command to an URI at [24] to get the acceleration in X, Y and Z axis. Thereafter, a translation algorithm is used to convert the sensor values onto mouse movements on the target device. By using Gesture Based Toolkits [21][22] mouse events like clicks and double clicks can be made just using Mouse Motion. This could be useful in Smart Space, wherein the users of the Smart Space needs to Control/navigate a big display screen by using their smart phone as a remote mouse. This application demonstrates that while using the RESTful Platform Broker, resources are not constrained onto a particular device and can be made available Smart Space wide. Further, Smart Space programmers can evolve the Smart Space by composing resources (e.g., Accelerometer) and making new services (Wireless Mouse), which were not earlier available.

Reading Contents of a File using a Remote Text To Speech (TTS): Most of the commercially available Android Platforms contains a Text to Speech Engine. This TTS can be made available across the Smart Space through the RESTful Platform Broker and can be used by other devices inside a Smart Space. This arrangement can be used to read a content of file which is present on a remote system (e.g., Linux based Laptop) through TTS present in the Android device. The RESTful Platform Broker exports the TTS at [29]. The Smart Space programmer can write a small program which issues a HTTP PUT command and send a payload consisting of an XML which contains the content to be spoken by the TTS in the Android. This way, the TTS is not restricted to the Android device but is available throughout the Smart Space.

Recording and Playback of Meeting Notes: A common scenario in a smart meeting room is to record minutes of the meeting as soon as the meeting starts. Thereafter the minutes can be played back at a different point in time. RESTful Platform Broker can be used to realize this scenario. The RESTful Platform Broker, exports the Media Recorder at [30]. A very simple client application can be made on any platform to send a HTTP PUT command to start and stop the Recording. The stimulus to start/stop recording may be gathered by systems like RFID or simply by the meeting convener. Later on, the recorded information can be played back using the exported Media Player resource.

Location Aware Services: The RESTful Platform Broker can be used to make Location aware services. Android Based systems typically have a GPS. The RESTful Platform Broker can be used to share the location information from the GPS on the Android throught the Smart Space. The RESTful Platform Broker exports the GPS on the Android device as a URI at [31]. A client application can issue a HTTP GET command to read a geographic location sensed at a particular time (also called a ”fix”). The location information consists of latitude and longitude, a UTC timestamp and optionally, information on altitude, speed, and bearing. This information can then be used by other systems within the Smart Space to design and develop Location Based Application. For example, an application could be designed and developed on a client to do Google search based on the location provided by the GPS. Therefore, a case like searching a hospital through Google could lead to results which are nearer to

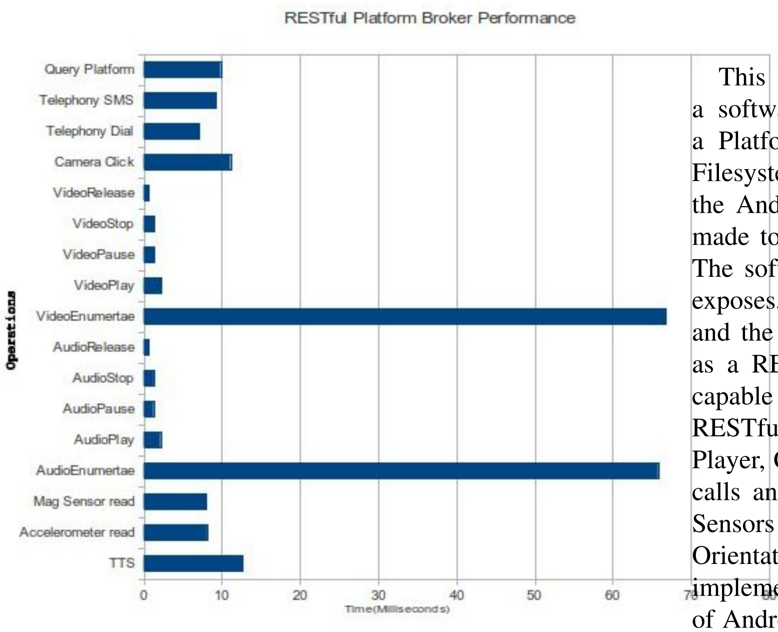


Figure 4. Performance of RESTful Platform Broker

a person who is in the Smart Space.

V. Application Profiling

The test setup for profiling RESTful Platform Broker includes a DELL Latitude E5400 Laptop machine, with Intel Core 2 Duo CPU T7250 at 2.00GHz, 1 GB RAM, 160 GB Western Digital Hard Disk Drive running GNU/Linux. This laptop interact with the RESTful Platform Broker installed on Samsung Based Android Phone (Samsung Galaxy Pop GT-S5570) using a Wireless LAN. In the GNU/Linux system Laptop, a shell script using curl [23] was developed to send HTTP Request to the RESTful Platform Broker on the Phone. While the HTTP Requests was being handled by the RESTful Platform, an application trace file is created in the sdcard of the Android Phone. Android SDKs Trace view utility was then used to analyze the created trace file. The Trace captures the time taken while processing the request apart from other information. A graph is prepared based on the Trace file created while the RESTful Platform Broker was processing the HTTP requests. The graph in Figure 4 shows the comparisons of the time taken to complete various operations by the RESTful Platform Broker. It can be seen that the slowest operations were enumerating media files. This may be attributed to the nature of this operation, which involved searching the Flash Drive for media files. Other operations do not require much of an I/O operation and are hence much quicker. We plan to further enhance the responsiveness of RESTful Platform broker with respect to these operations.

VI. Conclusion

This paper described the design and development of a software system capable of exposing the resources on a Platform (e.g., Android) as RESTful Service Oriented Filesystem. This software has been implemented on top of the Android Platform and an experimental setup has been made to inter-operate with other systems like GNU/Linux. The software is called the RESTful Platform Broker, as it exposes, the entire platform namely the software resources and the hardware resources onto the Smart Space network as a RESTful resource. The RESTful Platform Broker is capable of exposing almost the entire Android Platform as RESTful resources. The exported resources include Media Player, Camera, Telephony Stack (through which Telephone calls and SMSs can be sent), Text to Speech Engine and Sensors such as Accelerometer, Magnetic Field Sensors, Orientation Sensors, Proximity Sensor, etc. Though the implementation of the RESTful Platform Broker is on top of Android, it can be easily ported to other platforms which support a Java Virtual Machine.

RESTful Platform Broker has been designed as a framework for engineering programmable Smart Spaces. Therefore Smart Space programmers can evolve the Smart Space by introducing new services by composition of existing services. However, unlike previous attempts like Gaia [4], our approach does not require use of a particular programming language or use a particular API to do Smart Space programming. Though this feature is similar to PlanB Operating system, the advantage of our approach is that, since it is based on a popular virtual machine, namely the Java Virtual Machine, it can be installed on larger number of device and platforms as compared to PlanB. Further, PlanB was built on Desktop based system, whereas our approach has been to use a Mobile Platform, namely the Android Platform from the inception of the project. One of the major advantages of the RESTful Platform Broker is the use of Hypermedia. Through the use of Hypermedia, it is envisaged that there would be a greater decoupling between the RESTful Platform Broker and the HTTP Clients that uses it. This decoupling, therefore, allows RESTful Platform Broker to grow independently with respect to the clients that use it. The paper also describes some of the scenarios which are typical in a Smart Space environment that can be easily implemented using RESTful Platform Broker, with some amount of programming. It is however, envisaged that with the improvements in Visual Programming, Scripting Languages and Domain Specific Languages (DSL), the use of RESTful Platform Broker, may become easier and maybe used by non programmers also. The current version of the RESTful Platform Broker uses the MAC mechanism to control the visibility of resources within the Smart Space. In the future versions of RESTful Platform we plan to incorporate OpenID [32] and OAuth [33] for implementing

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

a more robust security mechanism.

REFERENCES

[1] M. Satyanarayanan, “Pervasive computing: Vision and challenges”,IEEE Personal Communications, vol. 8, pp. 10-17, 2001.

[2] S. Helal, W. Mann, H. El-Zabadani, J. King, Y. Kaddoura and E. Jansen, “The Gator Tech Smart House: A Programmable Pervasive Space”, Computer, vol. 38, March 2005 pp. 50-60.

[3] E. Hensberg, N. Evans and P. S. Marble, “Service oriented filesystems”, IBM Research Report, RC24788 (W0904-091), IBM Research Division Austin, IBM Research Zurich, 2009.

[4] M. Roman and R.H. Campbell, “Gaia: Enabling active spaces”, Proc. 9th Workshop on ACM SIGOPS European Workshop, ACM Press, New York, pp. 229-234, 2000.

[5] J. Al-Muhtadi, R. Campbell, A. Ranganathan, S. Chetan and M. Mickunas, “Olympus: A high-level programming model for pervasive computing”, Proc. 3rd IEEE Intl. Conf. on Pervasive Computing and Communications, pp. 7-16, March 2005.

[6] J. Al-Muhtadi, R. Campbell, E. Chan, and J. Bresler, “Gaia Microserver: An Extendable Mobile Middleware Platform”, Proc. 3rd IEEE Intl. Conf. on Pervasive Computing and Communications, pp. 309-313, March 2005.

[7] B. Johanson, A. Fox and T. Winograd, “The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms”, IEEE Pervasive Computing 1(2), pp.67-74, April-June 2002.

[8] A. Fox, B. Johanson, P. Hanrahan and T. Winograd, “Integrating Information Appliances into an Interactive Workspace”, IEEE Computer Graphics & Applications, pp.54-65, May-June 2000.

[9] S. Sen, J. Erman and A. Gerber, “HTTP in the Home: It is not just about PCs”, Proc. ACM SIGCOMM Workshop on Home Networks (HomeNets), pp.90-96, 2010.

[10] J. Mathrick, J. Humble, C. Greenhalgh, S. Izadi and I. Taylor, “ECT: A Toolkit to Support Rapid Construction of Ubicomp Environments”, Proc. UbiComp 2004, pp.207-234, Springer, 2004.

[11] F. J. Ballesteros, E. Soriano, G. Guardiola, K. Leal, “Plan B: Using Files Instead of middleware abstractions for pervasive computing environments”, IEEE Pervasive Computing 6(3) pp. 58-65, Aug 2007.

[12] R. Pike, D. Presotto, S. Dorward, B. Flandrena, K. Thompson, H. Trickey, and P. Winterbottom, “Plan 9 from bell labs”, Computing Systems, vol. Vol. 8, pp.221-254, Summer 1995.

[13] R Pike, D. Presotto, K. Thompson, H. Trickey and P. Winterbottom, The Use of Name Spaces in Plan 9“, ACM SIGOPS Operating Systems Review, Vol.27(2), pp.72-76, Apr. 1993.

[14] E.V. Hensbergen and R. Minnich, ”Grave robbers from outer space: Using 9P2000 under linux“, USENIX 2005 Annual Technical Conference, FREENIX Track, pp.83-94, 2005.

[15] C. Prehofer, J. Gurp, V. Stirbu, S. Sathish, P.P. Liimatainen, C. Flora and S. Tarkoma, ”Practical web- based Smart Spaces“, IEEE Pervasive, Vol 9(3), pp.72-80, 2010.

[16] L. Richardson, ”Justice will take us millions of intricate moves“, http://www.crummy.com/writing/speaking/2008-QCon/, QCon San Francisco, 2008 [retrieved: March, 2012].

[17] R. Gossweiler, C. McDonough, J. Lin, and R. Want, Argos: Building a web-centric application platform on top of android, IEEE Pervasive Computing, vol.10(4) pp.10-14, 2011.

[18] R. Fielding, Architectural Styles and the Design of Network based Software Architectures, PhD thesis, Doctoral dissertation, University of California, Irvine, 2000

[19] E. Gamma, R. Helm, R. Johnson and J. Vlissides, Design Patterns: Elements of Reusable Object-Oriented Software, Addison- Wesley, 1994.

[20] Dns service discovery, http://www.dns-sd.org/ [retrieved: March, 2012].

[21] Easystroke: A gesture recognition application for x11, http://sourceforge.net/apps/trac/easystroke/wiki, [retrieved: March, 2012].

[22] Gestikk - Mouse gesture recognition in Ubuntu, https://launchpad.net/gestikk, [retrieved: March, 2012].

[23] Curl and libcurl, http://curl.haxx.se/, [retrieved: March, 2012].

[24] http://ipAddressOfAndroid/Sensors/Accelerometer [retrieved: March, 2012]

[25] http://ipAddress/mediaPlayer/songList/ [retrieved: March, 2012]

[26] http://ipAddress/mediaPlayer [retrieved: March, 2012]

[27] http://ipAddress/mediaPlayer/songName1 [retrieved: March, 2012]

[28] http://GovindPhone.local/camera [retrieved: March, 2012]

[29] http://ipAddress/tts [retrieved: March, 2012]

[30] http://ipAddress/MediaRecord [retrieved: March, 2012]

[31] http://ipAddress/GPS [retrieved: March, 2012]

[32] B. Ferg et al., OpenID Authentication 2.0Final, OpenID Community, Dec. 2007; http://openid.net/specsopenid-authentication-2_0.html [retrieved: March, 2012]

[33] M. Atwood et al., OAuth Core 1.0, OAuth Core Workgroup, Dec. 2007; http://oauth.net/core1.0. [retrieved: March, 2012]

Researching Motivational Factors Towards a Sustainable Electricity Consumption

Louise Jensen
IT University of Copenhagen
Rued Langgaardsvej 7
Copenhagen S, Denmark
loje@itu.dk

Line Mulvad
IT University of Copenhagen
Rued Langgaardsvej 7
Copenhagen S, Denmark
limu@itu.dk

Sofie Hauge Katan
IT University of Copenhagen
Rued Langgaardsvej 7
Copenhagen S, Denmark
shka@itu.dk

Nicolaj Valberg Sundahl
IT University of Copenhagen
Rued Langgaardsvej 7
Copenhagen S, Denmark
nvsu@itu.dk

Abstract— **In relation to Smart Grid, which is currently in the making, this article presents findings from a qualitative study researching the motivational factors towards a sustainable electricity consumption. While there is a tendency mainly to look at the technical aspects of the implementation, we wish to take a critical look at the current implementation issues from the users' point of view. Previous empirical studies indicate that social interaction and visualisation of real time energy consumption patterns can trigger more ecologically responsible behaviour. This paper focuses on exploring this assertion through a qualitative study of a design called the "The Social Electricity Meter" by revealing an indication of motivational factors to change ones electricity use based on social stimuli. By reflecting theoretically on how this kind of empirical data is essential when designing future Smart Grid experiences, we also evaluate the capability of the approach *Research Through Design* to gather insights about future social practice.**

Keywords - ***Smart Grid; Research Through Design; Eco-visualisation; social interaction; electricity consumption.***

I. INTRODUCTION

By the year 2050, the Danish Government is striving to make Denmark almost independent of fossil fuels estimating that up to 80 percent of produced energy will come from renewable sources such as wind, wave, and solar production. As the current energy infrastructure is not constructed to handle these kinds of energy sources, it requires major changes in our existing energy infrastructure for these desired outcomes to be actualised [1]. Several initiatives have been taken to develop this new infrastructure, which is widely referred to as Smart Grid – a yet undefined concept with the purpose of creating a more flexible electricity grid through intelligent communication between the load of energy in the grid, the consumers, and their devices. Because renewable energy sources are uncontrollable, the electricity infrastructure must adapt and support flexible consumption in order to effectively make use of the renewable energy.

Right now, the consumption in Denmark is concentrated around electricity peak hours. This requires power stations to produce larger amounts of energy at certain hours of the day. It is not currently possible to store electricity extracted from renewable energy. Therefore, to utilise the renewable electricity it has to be used immediately in order for it not to be wasted. To eliminate the waste the aim is to move the current consumption in peak hours to periods when there is an overload of electricity available in the grid [1]. Due to the many aspects and the complexity of Smart Grid as a concept, we focus solely on private electricity consumption and the problems of peaks in electricity consumption.

Company stakeholders (GreenWave, HP, IBM, and DONG) in the field of Smart Grid have tested Smart Meters with the purpose of providing visualised information to inform individual users. The individual users' specific energy usage is visualised in an attempt to foster awareness of one's own energy consumption and as an incentive to change one's consumption practices [2]. Smart Meter is a development of the regular electricity meter that uses visual models and representations to display. In this regard, Tiffany Holmes, professor in art and technology, appropriately uses the term *eco-visualisations* to describe real time dynamic data visualisations of energy consumption as a method for inspiring environmental stewardship [3].

Steen Kramer Jensen, official energy consultant at Energinet.dk, points out that incentives for intelligent electricity consumption cannot only be economical [1]. Thus it is necessary to explore what other possible incentives might persuade users to modify their electricity consumption patterns. Research undertaken by different company stakeholders within the field of Smart Grid identifies social relations as having an influence on change in energy consumption behaviour [1, 4]. Therefore, we wish to explore and discuss how social interaction and eco-visualisation influence the user's practice around electricity consumption patterns. This study aims to advance the understanding of motivational factors behind behavioural change, and should be seen as an early exploration that provides inspiration and foundation for future research.

II. Theory

Research has been undertaken in relation to Smart Grid, including the forementioned Smart Meters. Common to all this research is a focus on the technological aspects, or testing of new technologies. As professor of sociology, Elisabeth Shove argues: there is a tendency ”... to pay greater attention to the invention and acquisition of the new things than to the way such novelties are subsequently deployed in practice” [5, p. 2]. Thus research within the field of Smart Grid seems to primarily focus on how to implement new technologies that will affect the everyday life but with limited understanding of the actual users. Hence, the research appears not to take into account how technologies and practices are mutually influenced because they are affected by the meso and macro level context in, which they are located or deployed [5].

Johan Redström, docent in interaction design, argues that design is defined by the user, in the situated use, and shaped by the context and practices of the user. By using the approach Research Through Design we want to make use of the process in, which the user assigns meaning and function to the design in relation to their existing practices and experiences [6]. Fallman and Stolterman note that bringing forth an artifact to explore alternative designs is “shaped by the ambition to explore new solutions, new directions, new technology and new usage, to broaden the overall design space or to rock the boat, without necessarily trying to solve existing and well-defined problems” [7, p. 270]. Furthermore, the chosen approach to this study, is based on the belief that individuals create meaning based on their own “lived experiences of a concept or a phenomenon” [8, p. 57].

Smart Grid as a technology is still in the development stage. Research Through Design presents an opportunity to transform the world from a current state into a preferred state by introducing a design artifact. In this case, providing a way to explore the users experience with elements of technology design not yet invented [9]. Thus, our main approach is the presentation of the design for the user and the exploration of what and how the formation of meaning is created.

III. Design elements

To explore how social relations can have an influence on change in energy consumption behaviour we have designed what we call “The Social Electricity Meter” based on game and eco-visualisations theory. One of the purposes of the design is to inform users about their specific energy usage in an attempt to foster awareness of one’s own energy consumption patterns through visualisation and thereby create an incentive to change consumption habits [3].

By integrating game elements in our design we aim to explore how social interaction can influence our behaviour in relation to private electricity consumption and understanding of electricity. Game designer Chris Swain argues how games can be designed to affect social change: “Games are well-suited to communicating a shared understanding of a problem because they allow users to experiment with potential solutions in a safe setting and generate their own mental frames for how it works” [10, p. 806]. The foundation for taking a playful approach to the design is that games set a frame, which people can relate to. Given that the implementation of Smart Grid is complex, and the understanding of electricity is abstract, it is beneficial to bring it into a game context where participants can draw upon previous experiences and thereby know how to play by the rules. By creating a design inspired by game theory we try to raise the energy usage issue without causing the user to get a bad conscience. By integrating social interaction, we aim to facilitate a *community of practice*, a term from the social anthropologist Etienne Wenge who has shown that by making it possible to share experiences and knowledge users build social capital and accumulate new knowledge and innovation [10].

Game theoreticians Zichermann & Cunningham emphasise that an important game mechanic, and one of the key corners of gamification, is feedback [11]. Based on this, the design contained a direct feedback in the form of points and comments triggered by the informants’ performance, which indicate whether or not they were heading in the right direction.

IV. The design

The design consists of a visual interface designed using HTML and PHP displayed on an iPad, which was placed in each of the informants’ home. The visual interface consists of 5 elements; 3 “score-bars” presenting individual scores as well as the average score through lightning graphics, a windmill, and a button (Fig. 1). The informants’ were asked to register every time they used certain products (washing machine, coffee machine etc.) by touching the button on the screen. The button was either green or red to symbolise whether it was a period with an abundance of electricity available in the grid or not. To further illustrate that there is an abundance of electricity in the grid (when the button is green) the windmill rotates as if moved by wind. Whenever the informants registered electricity used in the “green period” they gained a point and vice versa if it was registered in the “red period”. The main function of the design was to show the two informants their individual point score as well as the average amount of points of both informants. This makes both of them able to compare their energy consumption continuously.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

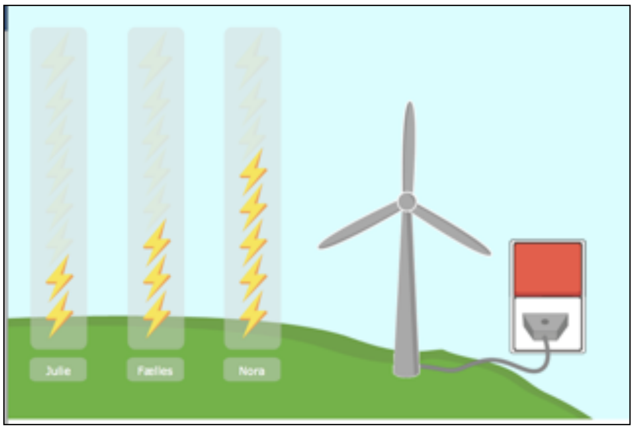


Figure 1. An example of the interface the user would see. In this case there is no abundance in the electricity grid, which is symbolized by the windmill not moving and the button turning red. If the user presses the button, which is now red, they will immediately get feedback in the sense of losing a "lightning-point"

Four informants, two pairs of friends, were using the design for seven days. The important selection criterion was to find pairs of informants who already had a social connection thus creating the best conditions for studying the impact of the social interaction. Two semi-structured qualitative interviews were afterwards conducted to follow up on their experience using the design, and how their electricity usage was influenced by the social aspect [12].

As we are not trying to provide categorical "truths", but rather raise questions of what social practices are possible, more informants would not necessarily have created more representative findings. Thus this study can be seen as exploratory by examining the use and formation of meaning in relation to a specific design [13]. The empirical data was then analysed by reducing the information to significant statements, and then placing it into themes [8].

V. FINDINGS

In our data analysis, five particular themes were identified:

A. Solidarity

Contrary to expectations, it turned out that the informants could not relate to the part of the design, which showed their common score. We expected that they would contact one another to cheer, motivate and tell when the energy was green with the aim of increasing their shared score. However, they expressed that they could not influence their partners' performance, thus giving no incentive to contact their partner. Additionally, they described a feeling of the common goal as being too abstract, which might be caused by the design not being distinct enough and giving no clear incentive to reach a common goal. This resulted in a lower support of direct social interaction and sense of solidarity between the participants than expected. Therefore, reaching the common goal of gathering points together was not a substantial part of their experience.

B. Discourse

An important finding was the indication of a strong established societal discourse constituting that one has to act environmentally responsible. This was distinctively visible in several ways but most significantly in verbal statements from all of the informants who implied that energy responsibility is common sense. It is interesting that the discourse was a motivational factor in the sense that it was the decisive term behind the desire to create a representation of oneself as environmentally responsible both in relation to others and to one self. Thus it was the major premise behind the following motivational factors.

C. Self presentation

It seems that it was a motivational factor for the informants that others were able to monitor their performance. The study indicates two different motivational aspects of self presentation in relation to electricity consumption. The first aspect is explicit as it is focused on actively presenting one's own results to others, e.g., by posting it on social network sites, or having the device a central place in the home. The other aspect is implicit as it is grounded in an automatic presentation of results, which is triggered by the system e.g., posting the electricity consumption of last week on a social network site.

D. Comprehension

In spite of the common goal being too abstract, the informants were unanimous in their expression that electricity consumption was made tangible and easier to comprehend when it was visualised through the design within a certain continuum. It was also considered motivating to see one's own performance compared with the friend's performance because this made their own consumption tangible.

E. Predictability

The informants indicated that they only had a limited amount of time and willingness to wait for the "green" electricity in order to use it. In certain instances, some of the informants mentioned that they did wait for the "green" electricity before performing a specific task. There would, however, be a larger willingness to wait if they could get an indication of when this was going to happen as suggested by one of the informants.

VI. DISCUSSION

Initially, we expected this study to show how social interaction would have an effect on energy consumption by creating a sense of community. On the contrary, what seemed to be a coherent tendency was the way social relations had an effect on the energy consumption in relation to the individual and self representation. This was due to either an implicit or explicit wish to represent one self as energy responsible. The sense of solidarity was not visible but instead the primary motivational factors for changing behaviour seemed to be on an individual level.

Research Through Design as an approach to explore future practices enabled us to zoom in on the detailed elements of electricity consumption and the behaviour around it. Thereby, we arrived at specific findings of the motivational factors in relation to electricity consumption.

Contrary to former research within the field of Smart Grid [1, 4] we have been able to explore the effect of social relations and also acquire information and an understanding of the underlying structures governing the motivational factors embedded in the social interaction. Because the design is put to actual use, existing practices are projected into the users' understanding of the design. Especially notable is how the aformentioned societal discourse on energy consumption has an influence on the use and understanding of the design. Most significant was how the elements of the design, which supports enrolment in the discourse, served as the motivational factors to change consumption behaviour.

The visualisation aspect of the design turned out to be a strong medium with the attribute of making the energy consumption tangible. Thus, it becomes a means for the user to create an understanding and a sense of self in relation to the above mentioned discourse on energy consumption. The pivotal factor for the visualisation aspect seemed to be the allocation of points reflecting energy responsible behaviour. It seemed, though, that the comparison of the points in relation to the other informant was what made the concept of consumption tangible. This was because of the ability to assign a value to the points on the basis of comparison. The value assigned to the points was primarily due to the social interaction through the design. Not only did they create a base for comparison with each other but also with one self over time. The comparison on different levels created a foundation for self representation. Thus, the motivational factor was both a clearer conscience and also representing oneself as an environmentally responsible person on the basis of the shared discourse.

Shove argues that the technology and the social discourses are coherently shaped by each other [5]. Thus, the understanding and the use of the design was very much assigned to this correlation. It also indicates that the design influences the discourse and the users' existing understandings of energy consumption. This seems evident in the way that the users, to some degree, have broadened their understanding of energy responsible behaviour from an understanding that energy responsible behaviour is only equivalent to reducing consumption, to a reformed understanding. A reformed understanding in, which it was not only about reducing consumption but also focused on what sources the energy derived from e.g., fossil fuels vs. renewable energy sources and how the production is distributed throughout the day.

VII. CONCLUSION

This study took starting point from other research within the field of Smart Grid, but with the method Research Through Design we took it into a more defined context and thereby gathered more detailed and focused data. An underlying societal discourse was located as the premise behind other motivational factors to change electricity consumption behaviours.

The visualisation enabled the user to compare their performance with others thereby making it more tangible and concrete as to what their contribution to a renewable environment looks like. Hereby transforming abstract and "diffused" information into comparable and comprehensible data and altering the understanding of what constitutes sustainable electricity consumption. Also, self representation played a part in the motivation for behavioural change. We have distinguished between whether the self presentation happens explicitly or implicitly through automated presentations.

The study has opened up the design space by showing how social interaction, because of the societal discourse can be a motivational factor towards sustainable energy consumption. The specific findings could be used as guidelines when researching and designing for behavioural change within the field of smart grid. Either as an inspirational point of departure, or a continuation of this research by redesigning "The Social Electricity meter" and going into depth with the elements, which were regarded as motivational.

REFERENCES

[1] Energinet.dk. Denmark opts for Smart Grid, "Intelligent power system with more renable energy." Magazine, Energinet.dk, 2011.

[2] Venables, M. "Smart meters make smart consumers." Engineering & Technology, 2007: 23-23.

[3] Holmes, T. "Eco-visualization: Combining Art and Technology." ACM, 2007: 153-162.

[4] Valocchi M., Juliano J., and Schurr A. "Lighting the way: Understanding the smart energy consumer." IBM Institute for Business Value, 2009.

[5] Shove, E., "Converging Conventions of Comfort, Cleanliness and Convenience." 2002: 1-16.

[6] Redström, J. "RE:Definitions of use." Design Studies, Elsevier Ltd., 29, nr. 4 (2008): 410-423.

[7] Fallman, D., and Stolterman E. "Establishing criteria of rigour and relevance in interaction design research." Digital Creativity, Routledge, 21, nr. 4 (2010): 265-272.

[8] Creswell, J. W., "Qualitative Inquiry & Research Design: Choosing among five approaches." London: Sage, 2007. ISBN 978-1-4129-1607-3.

[9] Zimmerman, J., Forlizzi J., and Evenson S. "Research Through Desing as a Method for Interaction Design Research in HCI." SIGCHI Conference on Human Factors in Computer Systems. San Jose, California: Carnegie Mellon University Research Showcase, 2007. 493-502.

[10] Swain, C. "Designing Games to Effect Social Change." DiGRA. DiGRA, 2007. 805-809.

[11] Zichermann, G., and Cunningham C. "Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps." 1st ed., Sebastopol, California.: O'Reilly Media Inc, 2011. p. 77. ISBN 978-1-449-39767-8.

[12] Kvale, S. "Interviews. An Introduction to Qualitative Research Interviewing." Thousand Oaks, California: Sage, 1996. ISBN 0-8039-5820-X.

[13] Silverman, D. "Doing Qualitative Research." 3rd ed., London: Sage, 2010. ISBN 978-1-84860-034-8.

Sensor Observation Service based Medical Instrument Integration

Data Capture, Analysis and Rendering

Chirabrata Bhaumik, Amit Kumar Agrawal, Suman Adak, Avik Ghose, Diptesh Das
TCS Innovation Lab
Tata Consultancy Services
Kolkata, India
e-mail: c.bhaumik@tcs.com, amitk.agrawal@tcs.com, suman.adak@tcs.com, avik.ghose@tcs.com, diptesh.das@tcs.com

***Abstract*—With advancement in technology, wireless portable medical instruments are quite common these days. However, these devices come with heterogeneous interfaces, which make them difficult to integrate the data generated in a common repository. Moreover, if a solution is created to integrate a few devices, adding a device with a new interface requires major effort and rework. This paper addresses this problem and suggests a method for integrating various health care instruments for use at home/small clinic on a generic Open Geospatial Consortium (OGC) Sensor Web Enablement (SWE) based Internet of Things (IoT) platform and exposing healthcare as a service. Finally the paper demonstrates how an example portal can be used to render the data on the server to end-users who may be patients, medical technicians or expert medical practitioners. This will enable governments or hospitals to offer a SMART Healthcare Services for citizens.**

Keywords- Smart Health; IoT; SOS; Healthcare; Mobile Application; Sensor networks; Intelligent Healthcare; Wireless Medical Devices; Smart HealthCare Analytics

I. INTRODUCTION

In developing countries like India, the general mass is not aware of the role of medical history in case management. Even if the medical history is kept, it is in the form of paper reports. These reports must be carried along while consulting with a healthcare professional. Also, these reports are at a higher risk of getting mutilated or lost compared to digital records. Hence, there is a need to digitize & store medical data in a central repository, which can be accessed by different parties (patients, home physicians, expert medical practitioners, medical institutions, insurance parties etc.) as and when required with proper access and authentication mechanism in place.

With the advancement in technology, wireless portable medical instruments are becoming reality. These devices are easy to operate, and can be used at home or small clinic with little or no training. These devices have the capability to communicate the medical observations to external world via digital interfaces. However, these devices come with heterogeneous interfaces and the data generated by them is in different formats. So, to store the data generated by these devices in a central repository, either manual intervention is required or a system must be there in place to take care of the idiosyncrasies of the varied interfaces & data.

Numerous relevant studies [3][4][5][7][8][9][10][16] have been found, out of which [3] and [4] are closer to the problem at hand:

Gavin E. Churcher [3] discussed the application of Sensor Web Enablement to a telecare application. It is a monitoring solution, which uses “dbFeeder” to directly insert data to SOS database. The major limitation of this work is not using the standard SOS interface to insert the observations.

M.V.M. Figueredo [4] explains a telemedicine system where the monitoring device communicates the medical data over RS-232 channel to a mobile phone, which in turn posts the data to a server through internet. Server stores the medical data in a custom designed database and presents the data as printed form or xml. The major limitation of this approach is the custom database schema, which necessitates a lot of rework to add a new monitoring device.

To address all these challenges, the authors have designed an Internet of Things (IoT) platform and offer healthcare as a vertical service to it. The IoT platform is generic enough to support other verticals as well. The bare minimum requirements for this IoT platform are as follows:

- Open Source implementation of Sensor Observation Service (SOS)[12] from 52North.org[2],
- Postgresql database with postgis plugin
- Proxy to handle REST requests from the clients and convert them to SOS compliant XML
- Tomcat servlet container

This platform provides a unified XML interface to communicate with the clients, and uses its own schema to store the data irrespective of incoming data formats. It eliminates the need to redesign the schema if data in a new format needs to be stored.

To provide healthcare service as an offering from this IoT platform, we implemented the following modules:

- A User Datagram Protocol (UDP) adapter
- A Bluetooth adapter

These adapters are small modules meant to bridge the gap between device interfaces (UDP and Bluetooth) and IoT platform interface (i.e. XML). We have chosen only these two types of adapters because we could get hold of wireless medical devices belonging to either of these two categories.

Section II describes the architecture of the whole system. In Section III, we describe the application components and

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

data flow. In Section IV, we provide the results and snapshots from our portal. Section V mentions the concluding remarks, and finally, Section VI acknowledges contributions made by various people in the organization to carry out this work.

II. SYSTEM OVERVIEW

Figure 1 illustrates the basic architecture of our proposed system. It is a layered architecture, which comprises of IoT platform, Plugin layer and Device Layer. Now, we shall take a closer look at each of these layers from the architectural perspective.

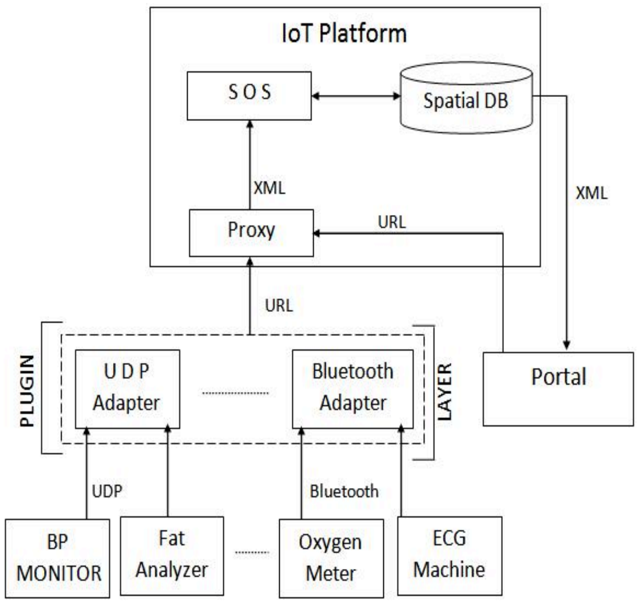


Figure 1. System Architecture

A. IoT Platform

The major offering from the IoT Platform is Sensor Observation Service (SOS). SOS is a part of the Sensor Web Enablement (SWE) [1][15] initiative from Open Geospatial Consortium (OGC). OGC standard states that SOS provides a generic model for all sensors and the observations collected by those sensors. It acts as a horizontal service, which can accommodate & store the observations from all types of sensors. It provides access to the sensor data in a standard way irrespective of the sensor type. SOS heavily depends on Observation & Measurement (O&M) standard [13][14] to model sensor observations. The SOS standard is described in full length in [11].

We are using Open Source implementation of SOS from 52North.org. It is a java web application and uses postgresql database to store the observations. It also supports spatial data types by using postgis plug-in from postgres. It accepts the sensor observations as O&M compliant XML document, and provides the response as O&M compliant XML document. This is the unifying ground on which observations from varied sensors flow into a common repository.

As O&M compliant XML documents are quite complex and resource consuming, it is difficult for low power devices to generate this kind of complex documents. Also it will involve network overhead, as only a few parameters change across the observations for a sensor. Hence, we have introduced a Proxy component in the IoT platform, which assumes the responsibility of generating the XML document and provides a parameterized Uniform Resource Locator (URL) as its interface. It simplifies the job of the sensors and reduces network overhead.

B. Plugin Layer

Wireless sensors do not follow any particular guidelines to provide an interface or observation format. The interface and data format completely depends on the manufacturer's discretion. The conventions used to send data & data format is significantly different for all types of devices.

So, it becomes necessary to provide a layer, which can listen on an interface supported by the sensor, parse the observation data and send it to the Proxy as a parameterized URL. We name this layer as Plugin layer. Whenever a new medical device is introduced with a different interface or data format, a module needs to be incorporated in this layer to support the interface and send the data to Proxy.

C. Device Layer

It is a virtual layer, which comprises of all the medical sensors. The diagram in Figure 1 depicts the devices that we have used in our work. Any new medical sensor will simply be added to this layer. Also, we are seeing another trend of new types of medical devices, which are coming out with a closed protocol. What we are observing that manufacturers of these devices are coming out with a portal of their own, and are also providing webservice Application Programming Interface (API) on request. So, the device layer can be easily plugged in as a module to get the medical data record from these portals as webservice API request.

III. APPLICATION DETAILS

We have worked with a few medical devices from ETComm such as Electrocardiogram (ECG) Machine [HC-201B], a FAT analyzer [HC-301W], a Blood Pressure (BP) monitor [HC-502W] & a Blood oxygen meter [HC801B]. These devices are neither 6LowPAN compliant, nor HL7 certified. So, we could not use these standards to communicate with the devices. All these devices map to the Device Layer as depicted in Figure 1.

Based on the communicative capability, these devices fall into two broad categories. The devices, which belong to the first category, communicate with a gateway over 433 MHz RF band. The gateway provides an interface to configure the IP address and port of UDP server where it can send the data. In our case, BP Monitor and FAT analyzer belongs to this category. The other categories of devices communicate with the external world via Bluetooth. In our case, ECG machine & Blood Oxygen Meter belongs to this category of devices.

The data formats for the observations generated by these devices are quite different. BP monitors and Fat analyzer

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

generates UDP payload where as ECG machine and Blood Oxygen meter generates a file.

To handle these interfaces, we have developed UDP adapter and Bluetooth Adapter modules on an Android phone. The UDP adapter module listens on an agreed upon port, identifies the device (both BP monitor and Fat analyzer), parses the data and sends them to proxy as parameterized URL. The Bluetooth adapter accepts the file, identifies the device, extracts and forwards data to proxy as parameterized URL. If the data consists of some binary field (e.g. image), it is sent as attachment to proxy layer and stored in IoT server either in the database, or in file system keeping a reference in IoT database. For instance, ECG Machine generates binary data, which is sent to proxy as attachment along with URL, whereas text data generated by Blood Oxygen meter is sent to proxy as parameterized URL. These modules map to the Plugin Layer as depicted in Figure 1. The medical devices used in this project do not provide alternate paths of communication. Each medical device provides just one interface and Plugin Layer provides listener for that interface (Bluetooth & UDP). XML is used as SOS interface which is part of the IoT platform. Figure 1 depicts the transmission path and role of these technologies in the whole application.

The parameterized URL contains a request string and a set of name value parameters. These request strings map to the SOS core operations. To store the observations, SOS provides two transaction operations:

- RegisterSensor - This operation registers a sensor with the SOS. It is required to register a sensor before storing any observation for it.
- InsertObservation - This operation is meant to insert an observation in the SOS repository.

To access the observations, SOS provides following operations:

- GetObservation - It is one of the mandatory core operation provided by SOS. It provides access to sensor observations based on the query. The query can involve both spatial as well as scalar data.

These operations are based on XML interface. The template XML conforming to SOS standard must be created for each combination of device types and SOS operation. All the sensor specific details and observation parameters must be carefully captured and mapped to SOS terminology. The following terms related to SOS demand a brief introduction as they are used extensively in our template XML documents:

- Sampling Time: The time when observation was made.
- Procedure: It is usually a sensor, which generates the observation. We have mapped the device identifier to procedure.
- Observed Property: These are the parameters that constitute the observation.
- Phenomenon: This is a characteristic parameter of the observation.
- Composite Phenomenon: It is sum total of characteristic phenomenon.
- Feature of Interest: It refers to the subject to which the observations relate to.
- Result: It is the value of a parameter created by a procedure.
- Observation: It is a process of observing.

The snapshot in Figure 2 depicts the XML we had designed for GetObservation Service for Blood Pressure.

```
<GetObservation xmlns="http://www.opengis.net/sos/1.0"
  . . . . .
  service="SOS" version="1.0.0"
  srsName="urn:ogc:def:crs:EPSG::4326">
  <offering>BLOOD_PRESSURE</offering>
  <observedProperty>
    urn:ogc:def:phenomenon:OGC:1.0.30:systolic
  </observedProperty>
  <observedProperty>
    urn:ogc:def:phenomenon:OGC:1.0.30:diastolic
  </observedProperty>
  <observedProperty>
    urn:ogc:def:phenomenon:OGC:1.0.30:pulsepressure
  </observedProperty>
  <observedProperty>
    urn:ogc:def:phenomenon:OGC:1.0.30:pulserate
  </observedProperty>
  <responseFormat>
    text/xml;subtype=&quot;om/1.0.0&quot;
  </responseFormat>
</GetObservation>
```

Figure 2. Request XML for BP GetObservation

Now, we take examples of Blood Pressure monitor, and Fat analyzer and illustrate the mapping to XML. As BP Monitor is the sensor, its unique device identifier can be used as procedure (or sensor Id). The four parameters viz. “Systolic Pressure”, “Diastolic Pressure”, “Pulse Rate” and “Pulse Pressure” measured by BP monitor maps to four Phenomena. The group of these four phenomena is mapped to Observed Property & Composite Phenomenon. The observation belongs to the patient, and so we have mapped the patient identifier as Feature of Interest. The result is mapped to the actual value for all the four parameters as reported by the BP monitor. For Blood Pressure, we have mapped the group of parameters as observed property.

For Fat analyzer, its unique device identifier can be used as procedure (or sensor Id). The Phenomenon will be changed to the five parameters measured by it, and accordingly the Observed Property & Composite Phenomenon will change. As, the observation belongs to the patient, so the mapping of Feature Of Interest will remain same as BP monitor. The Result is mapped to the actual value for all the five parameters as reported by the Fat analyzer. These mappings enable the creation of XML templates.

On receiving the parameterized URL, proxy component picks an XML template meant for the request string and device type, puts the parameters in the placeholders and sends it to SOS via HTTP POST.

The response generated by SOS depends on the SOS operation. On success, RegisterSensor returns AssignedSensorId whereas InsertObservation returns observation id. On failure, an exception or error message is issued.

It is evident from the above discussion that the medical instruments (even though provided by the same vendor) do not follow any guidelines to support uniform interface across all its devices and the same is true for the data format. Hence, we needed the plugin layer to handle these differences.

The URL requests originated from Plugin Layer map to the SOS transaction operations (as the sensors only store the observation, they don't fetch it). The URL requests originated from Portal (as depicted in Figure 1) maps to the GetObservation. The snapshot in Figure 3 depicts a typical XML response from SOS.


```
<om:Observation gml:id="go_1332167738541">
 . . . . .
 <om:procedure xlink:href="25155"/>
 <om:observedProperty>
  <swe:CompositePhenomenon
           gml:id="cpid2"
           dimension="5">
   <gml:name>Result Components</gml:name>
   . . . . .
   <swe:component xlink:href=
    "urn:ogc:def:phenomenon:OGC:1.0.30:pulserate"/>
   <swe:component xlink:href=
    "urn:ogc:def:phenomenon:OGC:1.0.30:systolic"/>
   <swe:component xlink:href=
    "urn:ogc:def:phenomenon:OGC:1.0.30:diastolic"/>
   <swe:component xlink:href=
    "urn:ogc:def:phenomenon:OGC:1.0.30:pulsepressure"/>
 </om:observedProperty>
 . . . . .
 <om:result>
  <swe:DataArray>
   . . . . .
   <swe:encoding>
    <swe:TextBlock decimalSeparator="."
          tokenSeparator="," blockSeparator=";"/>
   </swe:encoding>
   <swe:values>
    2011-12-12T13:08:48+05:30,9830576102,79.0,143.0,105.0,38.0;
    2011-12-12T13:10:44+05:30,9830576102,72.0,121.0,72.0,49.0;
    2011-12-12T14:32:07+05:30,9830576102,97.0,146.0,67.0,79.0;
    2011-12-12T15:39:33+05:30,9830576102,81.0,135.0,84.0,51.0;
    2011-12-12T19:08:08+05:30,9830576102,76.0,148.0,62.0,86.0;
    2011-12-13T14:29:15+05:30,9830576102,99.0,124.0,60.0,64.0;
   </swe:values>
  </swe:DataArray>
 </om:result>
</om:Observation>
```


Figure 3. Response XML from SOS containing BP Data

The Portal is the reporting tool that retrieves data from SOS using GetObservation operation. GetObservation provides querying capability based on the parameters like feature of interest (i.e., patient id), procedure (i.e., sensor id), offering, observed property, sampling time etc. To access the observations in SOS, a request is send to the proxy as parameterized URL. Proxy parses the URL and based on the request string, it turns generates the XML and send it to SOS. The response to this request is an XML with all the values meeting the query criteria. This XML is forwarded to the portal, where it gets rendered in tabular format.

Section IV describes the various medical devices we have worked with, the android code we had to develop for getting medical device readings via Bluetooth and the portal view of the data that has been kept in the central database using SOS.

IV. RESULTS

The Fat Analyzer and the BP Monitor belong to first category of devices, which communicate via vendor provided gateway and upload the medical sensor data via SOS using our UDP Adapter. Here, only the gateway needs to be configured for proper server IP and port for data posting. Once the data gets posted, the user can have a view of the medical readings in the portal. The snapshot in Figure 4 is depicting a typical portal view of Blood Pressure data.

However, for the second class of devices, where medical sensor data has to be captured via Bluetooth, we had to develop a mobile application to capture readings via Bluetooth and posting it to central database using SOS. The snapshot in Figure 5 depicts the landing page of our app, which waits for user input to take either ECG reading or Oximeter reading.

BP Measurement Data

Measurement_Time	User_ID	Systolic blood pressure	Diastolic blood pressure	Pulse pressure	Pulse rate
2011-12-13 14:29:15	9830576102	124.0	60.0	64.0	99.0
2011-12-12 19:08:08	9830576102	148.0	62.0	86.0	76.0
2011-12-12 15:39:33	9830576102	135.0	84.0	51.0	81.0
2011-12-12 14:32:07	9830576102	146.0	67.0	79.0	97.0
2011-12-12 13:10:44	9830576102	121.0	72.0	49.0	72.0

Figure 4. Portal View of BP Data

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

In the series of snapshots in Figure 8, we have depicted the steps one need to perform in our android app for taking the ECG data from the device, and getting it posted to the central database using SOS.

Figure 5. Landing Screen for the Android App

The second class of devices also need to have a local viewer application in the mobile as the user can check the readings and then push it to database once verified. The snapshot in Figure 6 depicts how we have implemented the ECG viewing scheme in our mobile application on Android. Once the data gets posted in database, Figure 7 and Figure 9 show the snapshot views from our Portal instance.

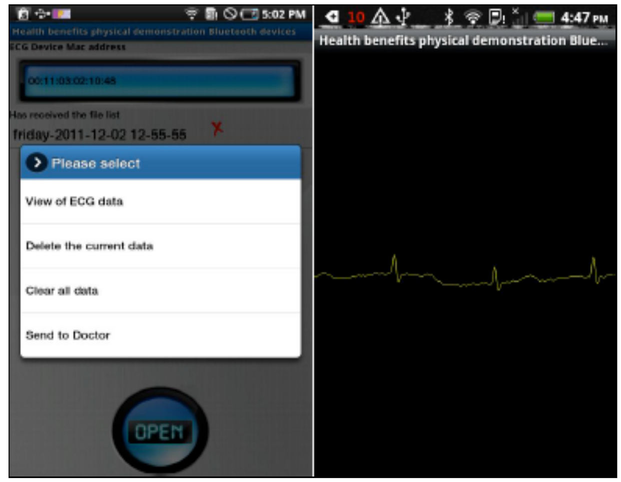


Figure 6. Local Viewing of ECG Data via Android App

ECG Measurement Data

Measurement Time	User ID	Heart Rate	Heart Behaviour	ECG Report
2012-01-12 14:10:05	patient1	76.0	NORMAL	Report
2011-12-16 11:21:59	pst	65.0	NORMAL	Report
2011-12-13 17:10:28	9830576102	67.0	NORMAL	Report
2011-12-13 17:08:52	9830576102	77.0	NORMAL	Report
2011-12-13 17:08:52	03366367415	77.0	NORMAL	Report

Figure 7. ECG Data Portal View

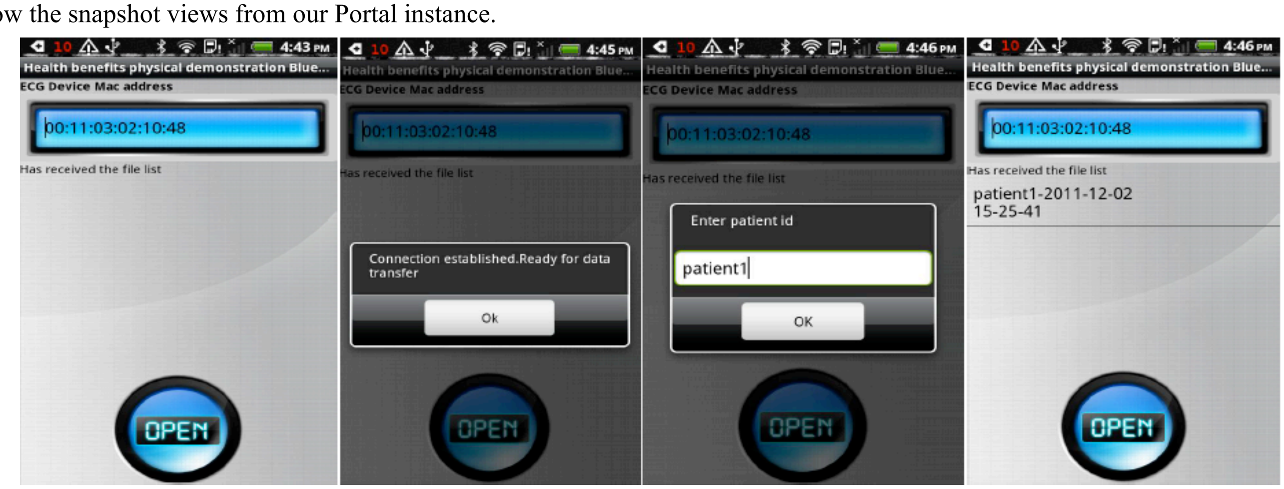


Figure 8. Steps for taking ECG via Android App

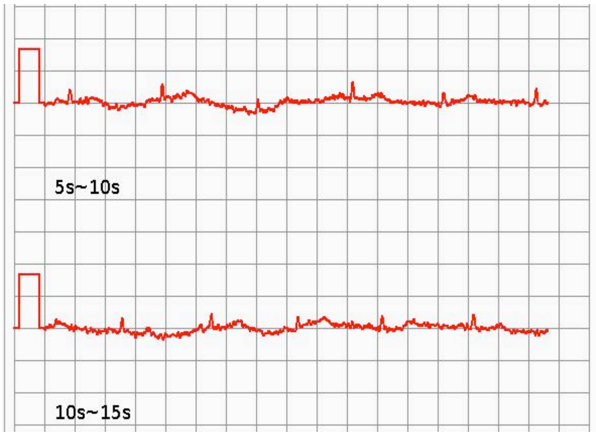


Figure 9. ECG Graph View from Portal

V. CONCLUSION AND FUTURE WORK

Our proposed application demonstrates healthcare as a vertical service to IoT platform. It also demonstrates the use of medical instruments with heterogeneous interfaces as web-enabled sensors, which contribute to a common SOS-enabled repository. The SOS layer enables the seamless integration of sensors by providing them a common ground, i.e. xml interface and SOS schema. Following are a few challenges and shortcomings of the system where improvement can be made:

- There exists devices which are closed for interaction and upload data directly to vendor's website for a portal view. Some of these vendors provide APIs to access this data from portal. The adapter, required at Plugin layer to incorporate these medical devices, needs to poll and gets these data as sensor observation into the IoT platform.
- The Healthcare service can really be given as a SMART service once we can bring SMART Analytics on the healthcare data. Currently we have implemented only the portal view of the healthcare data, but analytics services need to be run over this data to give more meaningful services to various stakeholders. So, the platform needs to be adapted for more analytics services to run over it.
- Mapping for medical streaming data to SOS compliant xml is not attempted in this application and can be taken as a potential future work.
- REST interface is not available in SOS implementations. While developing our application, we experienced that instead of XML interfaces, REST interface will be more efficient.
- Communication between Plugin Layer & Proxy as well as Proxy & Portal can be standardized using HL7.
- Available open source SOS implementations do not allow removal / editing of erroneous/corrupt data from SOS repository.
- The privacy and security of such sensitive data has not been taken care yet. Here, anybody with a credential can view anybody's medical data. This needs to be taken care in future version of the platform.

ACKNOWLEDGMENT

The authors would like to sincerely thank Mr. Arpan Pal and Mr. Prateep Misra for reviewing the paper. The authors would also like to thank Mr. Debnarayan Kar, Mr. Kallol Ghosh, Ms. Priyanka Sinha, Mr. Prasun Bhattercherjee and Mr. Joydip Das, all members of TCS Innovation Lab under the theme of Ubiquitous Computing for helping in implementation of the whole scheme for successful deployment of the project.

REFERENCES

[1] Botts M., Percivall G., Reed C., and Davidson J., "OGC Sensor Web Enablement: Overview and High Level Architecture," OGC 07-165, Version 3 Dec. 28, 2007, [retrieved: Apr., 2012], http://portal.opengeospatial.org/files/?artifact_id=25562

[2] Open Source SOS Software from 52North, [retrieved: Apr. , 2012], http://52north.org/communities/sensorweb/sos/

[3] Churcher G. E., Bilchev G., Foley J., Gedge R., and Mizutani T., "Experiences applying Sensor Web Enablement to a practical telecare application," Wireless Pervasive Computing, 2008. ISWPC 2008. 3rd International Symposium on, pp. 138-142, 7-9 May 2008

[4] Figueredo M. V. M. and Dias J. S., "Mobile Telemedicine System for Home Care and Patient Monitoring," Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE, vol. 2, pp. 3387- 3390, 1-5 Sept. 2004

[5] Sufi F., Fang Q., Mahmoud S. S., and Cosic I., "A Mobile Phone Based Intelligent Telemonitoring Platform," Medical Devices and Biosensors, 2006. 3rd IEEE/EMBS International Summer School on, pp. 101-104, 4-6 Sept. 2006

[6] Jirka S. and Bröring A., "Applying OGC Sensor Web Enablement to risk monitoring and disaster management," GSDI 11 World Conference, 2009, [retrieved: Apr., 2012] http://www.gsdi.org/gsdiconf/gsdi11/papers/pdf/96.pdf

[7] Varshney U., "Pervasive healthcare and wireless health monitoring," Mobile Networks and Applications, vol. 12, no. 2-3 (March 2007), pp. 113-127, DOI=10.1007/s11036-007-0017-1, http://dx.doi.org/10.1007/s11036-007-0017-1

[8] Sashima A., Inoue Y., Ikeda T., Yamashita T., Ohta M., and Kurumatani K., "Design and Implementation of Wireless Mobile Sensing Platform," SenSys'07, Nov. 6–9, 2007, Sydney, Australia, [retrieved: Apr., 2012] http://lecs.cs.ucla.edu/%7Enithya/SensysPSWorkshop/camera_ready/sashima.pdf

[9] Varshney U., "Pervasive Healthcare: Applications Challenges And Wireless Solutions," Communications of the Association for Information Systems, vol. 16(2005), art. 3, [retrieved: Apr., 2012] http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3018&context=cais

[10] Gupta V., Poursohi A., and Udupi P., "Sensor.Network: An open data exchange for the web of things," Pervasive Computing and Communications Workshops (PERCOM Workshops), 2010 8th IEEE International Conference on, pp. 753-755, March 29 2010 - April 2 2010

[11] McFerren G., Hohls D., Fleming G., and Sutton T., "Evaluating Sensor Observation Service implementations," Geoscience and Remote Sensing Symposium,2009 IEEE International, IGARSS 2009, vol. 5, pp. 363 –366, 12-17 July 2009

[12] Arthur N. and Priest M., "Sensor Observation Service," OGC 06-009r6, Oct. 2007, OpenGIS Implementation Standard, [retrieved: Apr., 2012], https://portal.opengeospatial.org/files/?artifact_id=26667

[13] Cox S., "Observations and Measurements – Part 1 - Observation schema," Version 1.0(2007a), OGC 07-022r1, Dec. 2007, OpenGIS Implementation Standard, [retrieved: Apr., 2012], https://portal.opengeospatial.org/files/?artifact_id=22466

[14] Cox S., "Observations and Measurements – Part 2 - Sampling Features," Version 1.0(2007b), OGC 07-002r3, Dec. 2007, OpenGIS

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

Implementation Standard, [retrieved: Apr., 2012], https://portal.opengeospatial.org/files/?artifact_id=22467

[15] Simonis I., "OGC Sensor Web Enablement Architecture," OGC 06-021r4, Aug. 2008, Version: 0.4.0, OpenGIS Best Practice, [retrieved: Apr., 2012], https://portal.opengeospatial.org/files/?artifact_id=29405

[16] Hashmi N., Myung D., Gaynor M., and Moulton S., "A sensor-based, web service-enabled, emergency medical response system," In Proceedings of the 2005 workshop on End-to-end, sense-and-respond systems, applications and services (EESR '05), USENIX Association, Berkeley, CA, USA, pp. 25-29

How to Co-Create Internet of Things-enabled Services for Smarter Cities

Sauro Vicini
eServices for Life and Health
Fondazione Centro S. Raffaele del Monte Tabor
Milano, Italy
e-mail: vicini.sauro@hsr.it

Sara Bellini
eServices for Life and Health
Fondazione Centro S. Raffaele del Monte Tabor
Milano, Italy
e-mail: bellini.sara@hsr.it

Alberto Sanna
eServices for Life and Health
Fondazione Centro S. Raffaele del Monte Tabor
Milano, Italy
e-mail: sanna.alberto@hsr.it

***Abstract* — What is the future of Smart Cities? The San Raffaele Scientific Institute (HSR)'s eServices for Life and Health unit in Milan strives to explore and push the boundaries of the Smart City concept through the ideation and implementation of smart services. Often, these services achieve their highest potential through Internet of Things, which enable the constituents of these services (users, products, environments) to be interconnected. In order to examine the dynamics between users, service touchpoints and Internet of Things, HSR decided to develop a methodology within a Living Lab framework and set up the City of the Future Living Lab. The City of the Future Living Lab is both a virtual and real research environment and community and embodies a Smart City (indeed it contains a university, laboratories, a hospital, offices, shops, a supermarket, post-offices, streets, parks, a light rail train and bus service, numerous ICTs, etc.). It therefore has exposure to all users and consumers of a city. This paper focuses on delivering an overview of the Living Lab methodology and the way it brings together people, environments, ICT and Internet of Things in the creation of e-Services designed by and around the end user. The paper presents the methodology and tools implemented for all the phases of the Living Lab process and presents the case of Living Labs as user-driven open innovation ecosystems for services for future Smarter Cities.**

Keywords - *Living Lab; Internet of Things; Smart City; Service Design; user-driven open-innovation*

I. INTRODUCTION

The concept of Smart City, and how it could evolve in the future, is a very hot topic. At the same time the advent of IoT (Internet of Things) as amplifier of traditional ICT (Information and Communication Technologies) has been recognized and is expected to greatly contribute to addressing today's societal challenges (from healthcare to energy efficiency, from education to well-being, from mobility to accessibility). In this context, the e-Services for Life and Health unit of San Raffaele Scientific Institute in Milan (http://www.eservices4life.org) explores, analyses, and develops tomorrow's services by fostering the collaboration between a number of different professionals, including doctors, engineers, designers, and psychologists. The aim of this research facility is to explore the ways in which the Smart City concept could evolve by involving users directly in the process for the creation of Internet of Things-enabled services within a real city environment. A user-driven open innovation ecosystem was adopted, successfully bridging the gap between fundamental and applied research. In order to integrate the strengths and engaging the efforts of users, stakeholders, researchers, SMEs (Small and Medium Enterprises) and policy makers, the City of the Future Living Lab was set up as a fruitful breeding ground for Smarter City Services.

The following paper will tackle the topic of the future of Smart Cities by firstly explaining how eServices for Life and Health interprets the concept of Smart City, and how Internet of Things and eService design combined can be considered essential elements for the development of services for Smarter Cities. The paper will then present the Living Lab methodology and process adopted by City of the Future Living Lab through descriptions of both, as well as through the presentation of five eServices that are being developed and implemented by the research unit as substantiation of the integration of Internet of Things and the Living Lab process.

II. SMART CITY & INTELLIGENT CITY

There is an array of terms used to define what is meant by Smart City in contemporary literature. In order to identify an exhaustive definition that successfully encompasses in a single concept all the most important elements, an evaluation of existing texts was made. The research was further broadened by the unit's team to include also definitions concerning Intelligent City in the search for the most successful umbrella term.

The report published by the Centre of Regional Science at Vienna University of Technology in 2007 [1] offers a holistic definition of Smart City, which encompasses a series of very valuable characteristics that the HSR find equally important. For a city to be considered a Smart City, it must call for the cooperation of a multitude of fields of activities including industry, education, community participation, technical infrastructure, and various 'soft factors': "A Smart City is a city well performing in a forward-looking way in six characteristics (Smart Economy, Smart People, Smart Governance, Smart Mobility, Smart Environment, Smart Living), built on the smart combination of endowments and activities of self-decisive, independent and aware citizens."

If analyzed, the above-mentioned definition can be broken down into the following characteristics. A Smart City must have well-developed connectivity obtained through a networked infrastructure [2] [3]. High-tech and creative industries (known as 'soft infrastructure') [4] must emerge from this fertile foundation and over time attract new businesses and investments therefore producing both urban growth and a positive socio-economic performance [5].

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

Concurrently, strong efforts must go into investing in social and relational capital, with the aim of creating a community that can successfully embrace technology and learn to comprehend, adapt and innovate it [6]. This in turn sets the base for social inclusion of urban residents within public services. Finally, for a city to be considered Smart, efforts must also go into investing in social and environmental sustainability, therefore successfully guaranteeing the safe and renewable use of natural heritage.

In reviewing the literature regarding the concept of Intelligent City, other factors previously not mentioned in the definitions regarding Smart City can be taken into consideration, which give good indications on the technology necessary in building a city environment that is well-connected, can be successfully enjoyed and used by inhabitants, is an inspiration for new business, helps nurture soft infrastructures, and is socially and environmentally sustainable. Intelligent Cities are often associated to the idea of a digital city, which unites a wide range of electronic and digital applications related to digital spaces of communities and cities [7] into an ICT network and uses this "to transform life and work within its region in significant and fundamental, rather than incremental, ways" [8]. These electronic and digital applications are interconnected among themselves via "embedded technology, which provides users with intelligent and contextually relevant support, augmenting our lives and our experience of the physical world in a benign and non-intrusive manner" [9]. The outcome of an Intelligent City, which for HSR is of extreme importance, is an innovation and ICT system that helps to combine the creativity of talented individuals, institutions that enhance learning and innovation, and virtual innovation spaces facilitating innovation and knowledge management [10], thus enabling "superior cognitive capabilities and creativity to be collectively constructed from combinations of individual cognitive skills and information systems that operate in the physical, institutional, and digital spaces of cities" [11].

For HSR, a Smart City is indeed a melting pot of all of the above features: a seamlessly interconnected ecosystem where products and environments interact among themselves across a number of ICT and Internet of Things-enabled services, with the objective of not only empower users, but also catalyzing the innovation process and bringing about positive social, economic and environmental change for governments, industries and citizens.

III. INTERNET OF THINGS AND ESERVICE DESIGN

In order to understand why Internet of Things and Service Design are interconnected, it is useful to identify the features of both elements. The idea at the heart of Internet of Things, which is very similar to the definition of Intelligent City, is that all things and all environments can be improved from a functional point of view via the embedding of technology that remains invisible to the eye of the users, which enables both products and environments to become smart: meaning that they can gather data (or enable someone to gather data via them) from their surroundings, producing what Fleisch calls high-resolution data (or real-time data, essential in management and improvement of systems [12]), as well as communicating amongst themselves as well as with humans. Internet of Things is a strongly evolving field and it is useful to keep in mind that in the future it will be open, scalable, flexible, secure, customizable by its users, profoundly user-centric [13].

Service Design is the activity of arranging and managing both tangible and intangible goods such as people, infrastructures, communication, documents and products, for the attainment of users' goals and the fulfillment of their needs. The aim of this process, as expressed by the Service Design Network Manifesto [14], is to create services that are useful, useable, desirable, efficient and effective, based on a human-centered and holistic approach that focuses on the customer experience whilst integrating team-based interdisciplinary approaches and methods, in ever-learning cycles.

The service economy in developing countries is increasingly accountable for higher percentages of GDP (Gross Domestic Product) and is mostly concentrated in financial services, health, and education. At the same time, products today have a higher service component than in previous decades, replacing the old dichotomy between product and service with a unified service-product continuum. This is leading to greater expectations and more articulated demands both from users (for increasingly smarter and engaging products and environments) as well as by service providers and infrastructures (for more streamlined, real-time and differentiating artifacts). For this reason, the role of Internet of Things is for it to become an integrated element of both services and their design process, for its capability of capturing, communicating and supplying data directly from the users and making it accessible to stakeholders and decision-makers.

Internet of Things can therefore be considered a powerful means through which to improve the value chain between users, products and environments and therefore better services. Providing adaptable, scalable and personalized services based on meaningful data collected via embedded technologies in products and environments pushes the concept of Smart Cities or Intelligent Cities towards a future where Cities are Smarter.

IV. TOWARDS EXPERIENCE DESIGN

Empowering users by delivering services that significantly improve their lives through augmented values and experiences is one of the key features of Intelligent Cities. At the same time, Smart Cities are those that stimulate communities to learn, adapt and innovate ICT and Internet of Things-enabled services, and therefore embrace technology. Thus it is safe to say that in both cases, users play a key role in defining the future of Smart Cities. Users' experiences can therefore constitute the principle driving force behind the development of successful and sustainable services for smarter cities. Disciplines such as Service Design [15], Transformation Design [16], Interaction Design [17] and Experience Design [18], exist and thrive by being cross contaminated by their users through a user-driven approach.

For these reasons, eServices for Life and Health, a department within HSR, has decided to adopt a user-driven approach to its innovation processes to produce an ecosystem where ideas are submitted to users' in order to gain their feedback, in a continuous process of open innovation [19]. The unit is specialized in the application of Information Technology to the realm of health and well-being, with the aim of developing and delivering services to the hospital's infrastructure as well as fostering innovation across numerous domains and disciplines. The reason for embracing a bottom-up approach is to have users contaminate the development process of an ICT or Internet of Things-enabled product or service so as to determine successful solutions that truly respond to (or even anticipate) user needs, bridging the gap between fundamental and applied research and diminishing the time and resources to deploy these on the market.

Such a process finds its natural collocation within a Living Lab methodology. In this context, City of the Future was brought to life, as a response to the unit's need to explore new practices to enrich ICT towards further innovation.

V. CITY OF THE FUTURE LIVING LAB

As previously anticipated, San Raffaele Scientific Institute has recently set up the City of the Future Living Lab, both a virtual as well as real research environment and community. The Living Lab follows along the conceptual framework presented by ENoLL (European Network of Living Labs) in which user-driven innovation is fully integrated within the co-creation process of new services, products and societal infrastructures.

The City of the Future Living Lab is an ecosystem where a multitude of stakeholders and partners can work alongside each other sharing knowledge whilst interacting with a wide variety of ICTs, therefore creating a fertile ground for disruptive innovation and cross-disciplinary research and communication.

A. *Living Lab Methodology*

The research methodology implemented by the City of the Future Living Lab is built on the widely recognized Living Lab process illustrated in Figure 1. The latter is based on four concurrent phases: co-creation, exploration, experimentation, and evaluation. Since this approach is an iterative and reflective one, a starting point is not defined and the LL process can be commenced at any stage of the design activity.

The fundamental concept at the base of the City of the Future Living Lab is to gain direct and unfiltered access to users' ideas, experiences, and knowledge, based on their daily needs and desire of feeling supported by products, services, or applications. Users are directly involved in co-creating, exploring, experimenting and evaluating new ideas, concepts and technological artifacts related to Internet of Things applications and services. The Living Lab methodology an approach that focuses on making stakeholders and users constructive and active participants in the definition and construction of an artifact, be it a product, an interface, a service or an Internet of Things-enabled platform, with the aim of improving and building value into these same artifacts. Users and stakeholders are all involved from the early stage of the LL process and throughout its entirety (therefore, along the Co-creation phase, the Exploration phase, the Experimentation phase, and finally the Evaluation phase as above). Insights are gathered directly from the users in order to define and implement realistic, useful, desirable and effective artifacts.

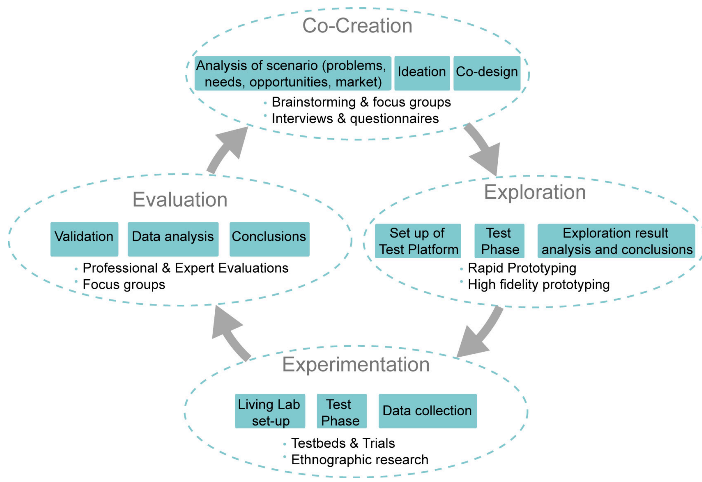


Figure 1: A visual representation of the Living Lab process

B. *Methods, Tools and Techniques used in the Living Lab Process*

The involvement of users in terms of an active integration of the end-users is a difficult task. One of the problems is that customer orientation is often based on classical market research tools, which puts customers in a passive "speaking only when spoken too" role [20]. City of the Future Living Lab studies, develops and implements a number of different tools throughout each phase of the living Lab process in order to best collect insights from users and use them to develop products and services that truly respond to their needs (both expressed and unexpressed).

The most commonly used tools during the co-creation phase of the Living Lab process are focus groups, interviews, brainstorming sessions and questionnaires. Interviews, questionnaires and focus groups are the most practiced exercises and have been widely discussed throughout academia [21] [22]. Focus groups consist in involving a group of users in a collective activity or discussion around their perceptions, opinions, beliefs and attitudes towards something that could be a product, ad advertisement, a packaging, and so on. This is a useful tool to gain a general overview of a target user as well as of a market. The same can be achieved by using interviews, which generally are one-to-one and involve an interviewer or researcher asking

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

open questions to an identified user. Questionnaires are most often composed of closed or semi-open questions though a moderator or researcher is not needed. All these tools are qualitative and can be used concurrently to build a clear picture of wither a market or a targeted user.

Another tool that can be used within the co-creation phase of the Living Lab process is the Lead User Method, developed by Von Hippel et al. [23] in the 1980s. It is a four-stage approach where user needs are explored by involving what Hippel calls "lead users", or expert users with extreme or very evolved requirements. Involving lead users at the birth of an idea allows researchers to delineate the contours of the market in which the idea could grow, benchmark existing solutions and therefore identify new opportunities.

Users can be effectively involved in the innovation process also via co-creation groups and idea competitions. A co-creation group is similar in structure to a focus group, but requires users to actively participate in the creation and development of an idea, be it a product or a service, throughout a group session. An idea competition instead is when users are asked to submit ideas to the research team within a specific time frame, and these ideas are then voted either by peers, stakeholders or a selected jury. In these two ways users are able to express their own creative interpretation regarding a new product or service, providing a fresh and often disruptive approach.

The previously mentioned tools are often used on a one-off basis and help to produce very focused and detailed insights. As users are becoming evermore active, informed and empowered, another tool that is very useful on a more continuous basis is consumer or user partnering. This consists of building a long-lasting relationship with a user who on a regular basis meets up with the research team and discusses his her lifestyle, habits, and behavioral evolutions. Crowd sourcing is another way of gaining insight over time. It can involve a group of users or single users, as well as experts or amateurs. A web-based platform or digital application can be created through which users are able to contribute their views and opinions on a given subject [24]. In this way, researchers and users co-create value through their continuous interaction towards a common goal.

An interesting emerging tool adopted during the co-creation phase is serious gaming. Serious games are indeed games (they can be digitally based or physical) but their aim is not only to entertain, but also to train individuals or investigate an issue. Different topics can be explored via serious games such as behaviors and attitudes in stressful situations, approaches to education, social dynamics, etc. Serious games are a dynamic and effective tool to help users access and use new information in an enjoyable manner, providing an enrichment of skills and living experience to its users. As with regular games, serious games fix a set of rules and each player has a designated role so that everyone is encouraged to participate. They successfully involve users of all social and cultural backgrounds and the "fun" aspect successfully triggers the spontaneous generation of a great number of insights.

In the exploration phase of the Living Lab process, the main tool used is prototyping. Prototypes can be made in two different manners: in a rough form (called low fidelity prototypes) and in a more finished yet not definite form (called high fidelity). Rough prototyping consists of making quick mock ups, and this exercise is used to simulate products and service components of an idea in order to better explain them to other members of the research team and stakeholders. High fidelity prototypes on the other hand consist of simulating a service's experience by setting up a representation of the service in mind via all its elements in a more finished form. Both ways are very pragmatic: the first approach allows to de-contextualize problems and issues and confront them with a blank or white slate manner, the second allows for the demonstration and testing of a solution and the drawing up of a set of requirements which will form the basis of the service which will be deployed in the experimentation phase.

Ethnography is a practice adopted from socio cultural anthropology and is based on a qualitative method where researchers examine the shared and learned patterns of values, behaviors, customers and beliefs in a group of users [25]. Researchers can observe users in their natural environment or in an artificially recreated one, though the prior is preferred because it allows users to feel at ease and therefore manifest more spontaneous attitudes. There are many techniques available, for instance, video ethnography, shadowing, disposable camera studies, day in the life studies, and so forth.

Ethnography is a very useful and commonly used tool during the experimentation of the Living Lab process. Through the study of users interacting with a product, service or environment in the process of being developed, the research team aims to understand the nature of behaviors between these elements. The insight gathered is fed back into the Living Lab process, forming the base for the evaluation phase, and allows the product, service or environment to be improved and therefore to respond more successfully to their users, often generating insight beyond the subject researchers initially intend to study.

The evaluation phase is the most critical step in the Living Lab process because it must identify the service's ability to permeate across users and their environment. The evaluation of a service can be viewed through different lenses – for example a business focus aims at understanding whether a service has the capacity of generating revenue over time or if it is sustainable. HSR's interest lies in understanding to what extent the services developed satisfy users' expectations and contribute in producing positive experiences. For this reason, the City of the Future Living Lab alongside a number of partners have recently implemented an experiential analysis platform as an outcome of an EU-funded project entitled ELLIOT (Experiential Living Labs for the Internet Of Things - http://www.elliot-project.eu/). Within this platform data are collected via the Internet of Things embedded in the products, services and environments around users (such as cameras, sensors, microphones, etc.), and provide a clear overview of their interactions and attitudes, thus providing researchers the opportunity to further optimize prototypes.

The Living Lab Process is explored and validated by applying a set of KSB (Knowledge-Social-Business) Experience Models to each service and correlated products developed. Such Models, specifically designed at the onset of each design process and based upon users' and stakeholders' objectives and expectations, evaluates whether the user and stakeholder objectives and expectations are met or not.

VI. EXAMPLES OF IOT SERVICES TESTED IN CITY OF THE FUTURE LIVING LAB

A. Interactive Totem Service

The Interactive Totem Service is based on a totem equipped with a touch screen monitor and an easy-to-use interface suitable for children as shown in Figure 2. It is placed in the pediatric department of HSR and allows hospitalized children to interact with a number of games that allow them to learn whilst playing. A reward system stimulates children to order their own meals and learn about nutrition as well as understand better how to manage their condition during hospitalization.

Figure 2: The Interactive TV Service within the hospital's pediatric ward.

B. Nutritionally Aware Vending Service

The vending service developed by HSR [26] offers a service for the promotion of healthy lifestyles, which stimulates users to reflect upon the benefits related to healthy living through a new concept of vending machines via an interactive touchscreen. The Vending Machine depicted in Figure 3 not only provides healthy foods, literature and music, as well as useful information regarding mobility and physical activity but also personalized information on how to adopt healthy lifestyle choices.

Figure 3: The Nutritionally Conscious Vending Service being developed within the Feed For Good project.

C. Bicycle Activity Monitoring and Tourism Service

This service is based on the Vainbici.it Web Portal and its aim is to promote initiatives for healthier and environmentally friendly lifestyles whilst providing new services for pedestrian-cycle mobility. It involves a digital platform (Figure 4) designed for exchanging and sharing information and digital content related to the world of bike and cycling and now includes a wearable monitoring system, which collects biological data from users as they cycles.

Figure 4: The Bicycle Activity Monitoring and Tourism Service involves a web platform that users can through different digital touchpoints.

D. Infomobility Service

The Infomobility Service proposed has the aim of improving the mobility of patients, visitors, staff and students in their travel to and from San Raffaele Hospital by offering a platform through which the latter can access a vast number of information including the timetable of the shuttle line, information about the condition of the automatic line and the next connections from and towards San Raffaele Hospital (Figure 5).

Figure 5: The app designed for the Infomobility Service involving the hospital's light rail shuttle line.

E. Energy Management Service

The Energy Management Service illustrated in Figure 6 has been developed as part of the eCube project, co-funded by the Italian Ministry of Economic Development and helps users monitor and improve their energy efficiency. Two pilots are being developed: an office pilot and a hospital pilot (which includes patients' rooms, corridors and nurses' room). Smart Appliances and Smart Plugs provided by project partners are being positioned in real work and hospital contexts and used by the users of these spaces.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

Figure 6: The energy consumption of an environment can be managed by the user in our Energy Management Service.

VII. Conclusions and Future Works

City of the Future Living Lab is a powerful representation of how user-driven open innovation ecosystems can contribute towards the evolution of the Smart City concept and the validation of Internet of Things-enabled services. In this real research environment and community that embodies an evolved vision of Smart City, a number of services have been co-created and will continue to be designed and offered across different scenarios, with the objective of studying and measuring the interactions between users and services as well as the potential of Internet of Things technologies and their impact on creating Smarter Cities.

The Living Lab process is actively and continuously contaminated by the user experience via a set of ever-evolving tools and techniques: during the co-creation phase, users can contribute to the ideation of services by suggesting their own ideas and experiences; user inspiration is explored by the research team through the development of prototypes; refined versions of the proposed services are then opened to users so that they can experiment with them and provide their very personal feedback; insights subsequently are used to evaluate the tested service and improve it.

HSR believes that this effective, user-driven process of innovation will play a vital role also in shaping the implementation of Future Internet technologies. As Smart Cities will embrace IoC-K (Internet of Content and Knowledge), IoP (Internet of People) and IoS (Internet of Services), Living Labs adopting a concurrently bottom-up and top-down approach achieved via a process of what we call co-creation, will become an evermore-important breeding ground for both incremental and especially disruptive innovation. This in turn will lead to the successful deployment of business, societal, economic and environmental change, which will provide the grounds upon which to build Smarter Cities.

Acknowledgment

City of the Future Living Lab would like to thank Rhea Vendors and S4Web for their technological contribution to the research projects mentioned.

References

[1] Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovic, N., and Meijers. E., “Smart cities – Ranking of European medium-sized cities”, Centre of Regional Science, Vienna (2007): http://www.smart-cities.eu/download/smart_cities_final_report.pdf (retrieved in January 2012)

[2] Hollands, R. G., “Will the real smart city please stand up?” City, Vol. 12, No. 3, pp. 303 - 320 (2008)

[3] Florida, R. L. “The rise of the creative class: and how it's transforming work, leisure, community and everyday life”, Basic Books, New York (2002)

[4] Domingue, J., Schaffers1, H., Komninos, N., Pallot, M., Trousse, B., Nilsson, M., and Oliveira, A., “Smart Cities and the Future Internet: Towards Cooperation Frameworks for Open Innovation”, Lecture Notes in Computer Science, Volume 6656, The Future Internet, pp. 431–446 (2011):

http://www.springerlink.com/content/h6v7x10n5w7hkj23/ (retrieved in Decembre 2011)

[5] Caragliu, A., Del Bo, C., and Nijkamp, P., “Smart cities in Europe”. Serie Research Memoranda 0048, VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics, pp. 131– 143 (2009)

[6] Coe, A., Paquet, G., and Roy, J. “E-governance and smart communities: a social learning challenge”, Social Science Computer Review, 19 (1), pp. 80-93 (2001)

[7] Droege, P., “Intelligent Environments - Spatial Aspect of the Information Revolution”, Elsevier, Oxford (1997)

[8] California Institute for Smart Communities, “Ten Steps to Becoming a Smart Community” (2001):

http://www.smartcommunities.org/library_10steps.htm (retrieved in January 2012)

[9] Steventon, A. and Wright, S., “Intelligent spaces: The application of pervasive ICT”, Springer, London (2006)

[10] Komninos, N., “The Architecture of Intelligent Cities”, conference Proceedings from the 2nd IET International Conference on Intelligent Environments, Athens, pp. 13-20 (2006)

[11] Komninos, N., “Intelligent Cities: Innovation, knowledge systems and digital spaces”, London and New York, Taylor and Francis, Spon Press (2002)

[12] Fleisch, E., “What is the Internet of Things? An Economic Perspective”, ITEM-HSG, Auto-ID Lab St. Gallen. (2006):

http://www.autoidlabs.org/uploads/media/AUTOIDLABS-WP-BIZAPP-53.pdf (retrieved in November 2011)

[13] Uckelmann, D., Harrison, M., and Michahelles. F., “Architecting the Internet of Things”, Berlin, Germany: Springer. ISBN 978-3-642-19156-5 (2011)

[14] Service Design Manifesto: http://www.service-design-network.org/system/files/media/sdn%20Manifesto.pdf (retrieved in November 2011)

[15] Parker, S. and Heapy, J., “The Journey to the Interface - How public service design can connect users to reform’. Demos, London (2006)

[16] Burns. C., Cottam, H., Vanstone, C., and Winhall, J., “RED PAPER 02 Transformation Design”, (2006): http://www.designcouncil.info/mt/RED/transformationdesign/TransformationDesignFinalDraft.pdf (retrieved in November 2011)

[17] Morridge, B., “Designing Interactions”, The MIT Press, Cambridge (2007)

[18] Norman, D.. “The Design of Everyday Things”, Basic Books, New York (1988)

[19] Chesbrough, H., “Open Innovation: The New Imperative for Creating and Profiting from Technology“, Harvard Business School Press (2003)

[20] von Hippel, E., “Successful Industrial Products from Customer Ideas”, Journal of Marketing, 42(1), pp. 39-49 (1978)

Copyright (c) IARIA, 2012.

[21] Mariampolski,. H., “Qualitative Market Research: A Comprehensive Guide”, Sage Publications, Thousand Oaks (2001)

[22] Kuniavsky, M., “Observing the User Experience: A Practitioner’s Guide to User Research”, Morgan Kaufmann (2003)

[23] von Hippel, E., “Lead Users: A Source of Novel Product Concepts”, Management Science, Vol. 32, No. 7. July, pp. 791-805 (1986)

[24] Howe, J. “The Rise of Crowsourcing”, Wired magazine, June edition (2006): http://www.wired.com/wired/archive/14.06/crowds.html (retrived in Januray 2012)

[25] Angrosino, M., “Doing Ethnographic and Observational Research”, Sage Publications, London (2007)

[26] Vicini, S., Sanna, A., and Bellini, S., “A Living Lab for Internet of Things Vending Machines”, Communications in Computer and Information Science, Volume 282, The Impact of Virtual, Remote, and Real Logistics Labs, Part 2, pp. 35-43 (2012): http://www.springerlink.com/content/j2t467q358l07

Wi-Fi Proximity as a Service

A new approach for distributing hyper-local data

Dmitry Namiot
Lomonosov Moscow State University
Faculty of Computational Mathematics and Cybernetics
Moscow, Russia
dnamiot@gmail.com

Manfred Sneps-Sneppe
Ventspils University College
Ventspils International Radio Astronomy Centre
Ventspils, Latvia
manfreds.sneps@gmail.com

Abstract—**This paper describes a new model for delivering hyper-local data for mobile subscribers. Our model uses Wi-Fi proximity as a service. In our concept, any exiting or even a specially created Wi-Fi hot spot could be used as presence sensor that can trigger access for some user-generated information snippets. In this approach we can discover hyper local data as data snippets that are valid (relevant) for mobile subscribers being at this moment nearby some Wi-Fi access point. And an especially developed mobile application (context-aware browser) can present that information to mobile subscribers. As the possible use-cases, we can mention for example news and deals delivery in malls, news feeds for office centers and campuses, Smart City projects, personal classifieds, etc.**

Keywords-Wi-Fi; proximity; collaborative location; indoor positioning; context-aware browsing.

I. Introduction

Technically, proximity sensor is the sensor able to detect the presence of nearby objects without any physical contact [1]. In this article, we will describe a model where the exiting or even especially created wireless networks could be used as proximity sensors. The detected proximity will trigger hyper-local news data delivery to mobile subscribers.

Many mobile applications can be characterized as collaborative in the sense that mobile nodes use the wireless network to interact with other mobile nodes that have come together at some common location. Collaborative nodes typically come together in some area, establish associations with other collaborative nodes dynamically, and make use of common services already available in that locality or provided by members of the group. The members of such a group may migrate together, like the visitors in the mall, some group of pedestrian, etc. [2].

Although these applications may use infrastructure networks, they will often use ad hoc networks since they are immediately deployable in arbitrary environments and support communication without the need for a separate infrastructure. This collaborative style of application may be useful in the ubiquitous computing [3]. Context awareness is defined as complementary element to location awareness. Whereas location may serve as a determinant for resident processes, context may be applied more flexibly with mobile computing with any moving entities, especially with bearers of smart communicators. Context awareness originated as a term from ubiquitous computing or as so-called pervasive computing which sought to deal with linking changes in the environment with computer systems, which are otherwise static [4].

Modern applications adopt a context-aware perspective to manage:

a) Communication among users and among systems, or between the system and the user,

b) Situation-awareness, like modeling location and environment aspects (physical situation) or the current user activity (personal situation)

c) Knowledge chunks: determining the set of situation-relevant information, services or behaviors [5].

In our approach, we are dealing with context-aware knowledge chunks. Let us start with the base element – location.

There are many different approaches for getting location info for mobile subscribes. In general, it could be pretty standard nowadays (GPS, cell-id, assisted GPS [6]), but everything is getting more complicated as soon as we need indoor positioning. Due to the signal attenuation caused by construction materials, the Global Positioning System (GPS) loses significant accuracy indoors. Instead of satellites, an indoor positioning system (IPS) relies on nearby anchors (nodes with a known position), which either actively locate tags or provide environmental context for devices to sense. The localized nature of an IPS has resulted in design fragmentation, with systems making use of various optical, radio, or even acoustic technologies [7].

Nowadays, a great number of technologies are being used for indoor localization, such as Wi-Fi, RFID etc. [7]. However, all of them require the utilization of their own API with their own protocols. This can be a big challenge for developing heterogeneous scenarios where different localization systems have to be used for a location service.

One of the most used approaches to indoor location is Wi-Fi based positioning. A standard Wi-Fi based positioning system, such as the one offered by Cisco is completely software-based and utilizes existing Wi-Fi access points installed in a facility and radio cards already present in the user devices. Companies could deploy also Wi-Fi based radio tags that use industry standard components that adhere to the 802.11 standards. This approach allows for the use of commercial off-the-shelf hardware and drivers to produce a

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

standards-based radio tag that can communicate bi-directionally over the 802.11 networks.

Thus, a standard Wi-Fi based positioning system can realize any type of location-aware application that involves PDAs, laptops, bar code scanners, voice-over-IP phones and other 802.11 enabled devices. For embedded solutions, there is no need for the client to include a specialized tag, transmitter, or receiver.

Because of the entire use of standards-based hardware, such as 802.11b, 802.11g, and 802.11a, a standard Wi-Fi based solution rides the installed based and economies of scale of the networks and end user devices that are proliferating today. Without the need for additional hardware, a company can install the system much faster and significantly reduce initial and long-term support costs. A common infrastructure supports both the data network and the positioning system, something companies strive for. The positioning system works wherever there is Wi-Fi coverage.

Wi-Fi location positioning is based on a grid of Wi-Fi hotspots providing, in general, 20–30 meters location accuracy. For more accuracy, there needs to be more access points. There are many articles devoted to Wi-Fi positioning. For example, we can combine a reference point-based approach with a trilateration-based one etc. Several layers of refinement are offered based on the knowledge of the topology and devices deployed. The more data are known, the better adapted to its area the positioning system can be [8].

Lets us mention also one more interesting approach: collaborative location (CL) [9]. The most interesting approach for our future development is Collaborative Location-sensing. Cooperative Location-sensing system (CLS) is an adaptive location-sensing system that enables devices to estimate their position in a self-organizing manner without the need for an extensive infrastructure or training.

Simply saying, hosts cooperate and share positioning information. CLS uses a grid representation that allows an easy incorporation of external information to improve the accuracy of the position estimation [9].

The motivation for CL and CLS is very transparent. In many situations, due to environmental, cost, maintenance, and other obstacles, the deployment of a dense infrastructure for location sensing is not feasible. It is exactly what we wrote about infrastructure-less system. In CLS, hosts estimate their distance from their neighboring peers. This can take place with any distance estimation method available (e.g., using signal strength). They can refine their estimations iteratively as they incorporate new positioning information.

Another interesting aspect for our approach is dynamic location based services. For example, AROUND [10] architecture is proposed as an approach for supporting location-based services in the Internet environment. AROUND provides a service location infrastructure that allows applications to select services that are specifically associated with their current location. The architecture includes a flexible scope model that defines the association between services and location, and a service location infrastructure organized by spatial criteria and optimized for location-based queries.

And at this point we are ready to make the last proposition before switching to our SpotEx model. Of course, the acronym LBS (Location Based Systems) contains the word “location”. But do we really need the location for the most of the services? As seems to us the final goal (at least for the majority of services) is to get data related to the location, rather than location itself. Location in the classical form (latitude, longitude) here is just an intermediate result we can use as key in our requests for obtaining data (our final goal). So, why do not request data directly if we can estimate location? We need to explain, what does it mean “estimate” here. It is exactly the proximity. But here are the obvious obstacles. The location services have got a wide support on mobile world for example. We should make proximity as easy to use for developers as location has now become. Then we will unleash a similar explosion of innovative services.

II. Related work

Lets us overview some existing projects. AllJoyn [11] is a peer-to-peer technology that enables ad hoc, proximity-based, device-to-device communication without the use of an intermediary server. Technically, it is a set of APIs for developers.

Nokia Instant Community is a new, instant way for communities to socially interact when in close proximity, without the need for WLAN infrastructure or Bluetooth and cellular connections [12]. The clever thing about Nokia Instant Community is that it doesn’t need the Internet. It means no searching for a Wi-Fi hotspot to get you in touch, and you don not need infrared or bluetooth either. It works completely by using the device’s adhoc Wi-Fi. But we should mention here that the whole idea looks very similar to Wi-Fi direct spec.

LocalSocial by Rococosoft [13] offers a proximity platform for sharing social data. LocalSocial combines information about people and things close to a user, with information from one or more of their social networks, and makes that information available to a mobile application in a simple way.

LocalSocial has 2 key components:

1. The LocalSocial Mobile Library, which is deployed on the handset. This gives developers APIs that allow them to:

a) Use proximity to detect devices nearby using Bluetooth, WiFi, and NFC

b) Connect securely to the LocalSocial Service

c) Discover information about the devices nearby - including opt-in social tags and information

e) Send messages to devices nearby, store tags about those devices for later

2. The LocalSocial Cloud Service is a web-hosted service, which mediates between the mobile application and the cloud, and stores and retrieves information about the devices (as tags), tracks analytics, and provides social connections to key social networks.

What are the common points to all projects? It is device based API. At the first hand, they are oriented to the device – to-device proximity. It is probably more convenient to the social networks. In our project, we will target device-to-infrastructure objects proximity. It is more convenient to the information providing. For example, deliver some news data for people in Smart City applications; provide dynamic information systems in campuses and office buildings, etc.

III. OUR APPROACH - SPOTEX

What if we stop our traditional indoor positioning schema on the first stage: detection of Wi-Fi networks? This detection actually already provides some information about the location – just due to local nature of Wi-Fi network. And as the second step we add the ability to describe some rules (if-then operators, or productions) related to the Wi-Fi access points. Our rules will simply use the fact that the particularly Wi-Fi network is detected. Based on this conclusion we will open (read – make them visible) some user-defined messages to mobile terminals. Actually, it is a typical example for the context aware computing. The visibility for user-defined text (content) depends on the network context.

The first time this service SpotEx (Spot Expert [14]) was described by the authors in article published in NGMAST-2011 proceedings [15]. It is a working application for Android platform.

Obviously, our SpotEx model is based on the ideas of Wi-Fi proximity. Any Wi-Fi hot spot works here just as presence sensor. But we are not going to connect mobile users to the detected networks and our suggestion does not touch security issues. It is not about connectivity at all. We need only SSID for networks and any other public information.

So, our service contains the following components:

- database (store) with productions (rules) associated with Wi-Fi networks. It is web-based data store
- rule editor. Web application (including mobile web) that lets users add (edit) rule-set, associated with some Wi-Fi network
- mobile applications, that can detect Wi-Fi networks, check the current conditions against the database and execute productions

How does it work? We can take any exiting Wi-Fi network (or networks especially created for this service – the most interesting case, see below) and add some rules (messages) to that network. Message here is just some text that should be delivered to the end-user's mobile terminal as soon as the above-mentioned network is getting detected via our mobile application. The word "delivered" here is a synonym for "available for reading/downloading".

The possible use cases, including commercial deployment are obvious. Some shop can deliver deals/discount/coupons right to mobile terminals as soon as the user is near some predefined point of sale. We can describe this feature as "automatic check-in" for example. Rather than directly (manually or via some API) set own presence at some place (e.g., similar to Foursquare, Facebook Places, etc.) and get deals info, with SpotEx mobile subscriber can pickup deals automatically. Campus admin can deliver news and special announces, hyper local news in Smart City projects could be tight (linked) to the public available networks and delivered information via that channel etc.

Especially, we would like to point attention to the most interesting (by our opinion, of course) use case: Wi-Fi hot spot being opened right on the mobile phone. Most of the modern smart phones let you open Wi-Fi hot spots. We can associate our rules to such hot spot (hot spots) and so our messages (data snippets) become linked to the phones. Practically, we are getting dynamic LBS here: phone itself could be moved and so, the available data will be de-facto moved too.

Figure 1. Wi-Fi hot spot on Android

This use case is probably the most transparent demonstration of SpotEx model. We can open "base" network right on the mobile phone, attach ("stick") rules for the content to that network and it is all do we need for creating a new information channel. There is no infrastructure except the smart phone and we do not need a grid of devices as per CLS models. By the way, it is the main difference from the centralized location frameworks from Goggle, Nokia or Ericsson, too. All the frameworks for Wi-Fi positioning are using static databases.

And note again that this approach does not touch security and connectivity issues. You do not need to connect mobile subscribers to your hot spot. SpotEx is all about using hot spot attributes for triggers that can discover the content. The term Wi-Fi proximity is used sometimes in connection with Wi-Fi marketing and mean on practice just setting a special splash screen for hot spot that can show some advertising/branded messages for users during the connection to that hot-spot. Unlike this, SpotEx threats Wi-Fi hot spots just as sensors.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

How our productions data store (base of rules) looks like? Each rule looks like a production (if-then operator). The conditional part includes the following objects:

Wi-Fi network identity,
signal strength (optionally),
time of the day (optionally),
client ID (see below).

In other words, it is a set of operators like:

IF *network_SSID* IS 'mycafe' AND *time is 1pm – 2pm* THEN { present the coupon for lunch }

Because our rules form the standard production rule based system, we can use old and well know algorithm like Rete [16] for the processing. A Rete-based expert system builds a network of nodes, where each node (except the root) corresponds to a pattern occurring in the left-hand-side (the condition part) of a rule. The path from the root node to a leaf node defines a complete rule left-hand-side. Each node has a memory of facts, which satisfy that pattern. This structure presents essentially a generalized tree. As new facts are asserted or modified, they propagate along the network, causing nodes to be annotated when that fact matches that pattern. When a fact or combination of facts causes all of the patterns for a given rule to be satisfied, a leaf node is reached, and the corresponding rule is triggered [17].

The current implementation for mobile client based on Android OS. This application uses *WiFiManager* from Android SDK - the primary API for managing all aspects of Wi-Fi connectivity. This API let us pickup the following information about nearby networks:

SSID - the network name.
BSSID - the address of the access point.
capabilities - describes the authentication, key management, and encryption schemes supported by the access point.
frequency - the frequency in MHz of the channel over which the client is communicating with the access point.
level - the detected signal level in dBm.

So, actually all the above-mentioned elements could be used in our productions. And now we can prepare rules like this:

IF *network_SSID* IS 'mycafe' AND level > -60db AND time is 1pm – 2pm AND network_SSID 'myStore' is not visible THEN {present the deals for dinner}

Block {*present the deals for dinner*} is some data (information) snippet presented in the rule. Each snippet has got a title (text) and some HTML content (it could be simply a link to external site for example). Snippets are presenting coupons/discounts info for malls, news data for campuses, etc.

Technically, any snipped could be presented as a link to some external web site/mobile portal or as a mobile web page created automatically by the rule editor included into SpotEx. Rule editor works in both desktop and mobile web. So, once again just having ordinary smart phone is enough for creating (opening) information channel for delivering hyper-local news data.

In case of presenting our data as links to some existing mobile sites (portals), SpotEx works as some universal discovery tool. De facto it lets mobile subscribers to be aware about context-relevant web resources. And owners for the mobile resources can describe own sites via rules rather then present individual QR-codes or NFC-tags for example.

In case of describing some content right in the SpotEx the whole system works in this part as content management system (CMS). SpotEx rule editor creates mobile web page for the each provided data snippet and hosts that page on the own server. It means by the way, that for presenting our data we can use any resources that could be presented on HTML pages. In particularly, any multimedia content is also supported.

SpotEx mobile application, being executed, creates dynamic HTML page from titles (according to rules that are relevant in the given context) and presents that mobile web page to the user. It works just as a classical rule based expert system: matches exiting rules against the exiting context and makes the conclusions. Existing content here is a description for "Wi-Fi environment": list of hot spots with attributes. And conclusion here is a list of titles that can be presented as a dynamically created mobile web page. On that page each discovered title could be presented as a hyperlink that points to the appropriate data snipped. Any click on the interested title opens the snippet (shows or discovers data to mobile user).

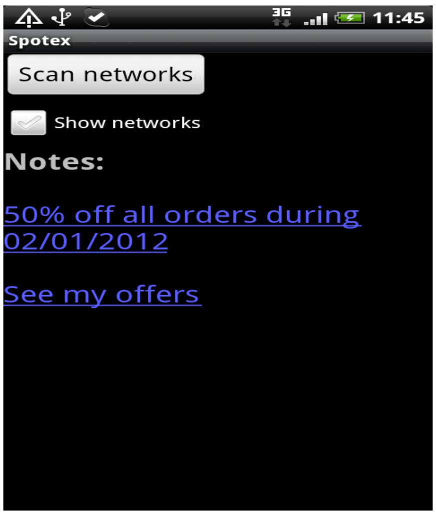


Figure 2. SpotEx client

So, for the mobile users the whole process looks like browsing, where their browser becomes aware about hyper-local content. It is a typical example of context-aware retrieval. Context-aware retrieval can be described as an extension of classical information retrieval that incorporates the contextual information into the retrieval process, with the aim of delivering information relevant to the users within their current context [18].

The context-aware retrieval model includes the following elements:

• a collection of discrete documents;
• a set of user's retrieval needs, captured in a query;
• a retrieval task, to deliver the documents that best match the current query, rated on the basis of a relevance measure;
• the user's context, used both in the query formulation and associated with the documents that are candidates for retrieval.

It is obviously, that all the above-mentioned tasks are components of SpotEx.

As per other functionality of our context-aware browser, we can highlight the following notes. At the first hand, we can note that it is the "pull model", versus the "push model" that proposed by Bluetooth marketing for example. It could be more convenient (more safe) for the users – there are no automatically downloaded files/messages, etc. But in the same time nothing prevents us from updating that dynamic web page automatically (e.g., by the timer) and simulating "pull model" in the user-safety mode.

At the second hand, we can note that because it is browsing, the whole process is anonymous. Indeed, there is no sign-in in the SpotEx. Of course, any data snippet may lead to some business web site/portal, where that site may ask about login etc., but the SpotEx itself is anonymous. Unlike social networks like Foursquare you do not need to disclose your identity just for looking mall's deals for example.

But, in the same time, we still can collect some meaningful statistics in SpotEx. Because the model requires Wi-Fi to be switched on, we have automatically unique ID for the each client. It is MAC-address. And it is actually global UUID. So, where we have not login info for our clients, we still can distinguish them. It let us detect for example, the same person, who did that already twice during the last week, opens that the particular data snipped.

Because mobile users in SpotEx model actually work with web pages, we can use pretty standard methods for web server log analysis for discovering user's activities.

A statistical analysis of the server log may be used to examine traffic patterns by time of day, day of week etc. So, we can detect frequent visitors, usage patterns etc. And even more – we can use that information in our rules. E.g., some mall may offer special things for frequent visitors, etc. Data from real time analytics for our info snippets could be used in conditional parts of our rules.

The next stage of development targets the simplicity of preparing data for SpotEx model. What if instead of the separate database with rules (as it is described above) we add the ability to provide a special markup for existing HTML files?

So, rather than writing separate if-then rules we can describe our rules right in HTML code. Technically, we can add for example HTML div blocks with attributes that describe our rules (their conditions). Now, using some JavaScript code we can loop over such div blocks and simply hide non-relevant from them. For doing that we need to make sure that our JavaScript code is aware about the current context. We can achieve that via a special light implementation of local web server. This web server, being hosted right on the mobile phone (on the Android in our case) responds actually only to one type of requests. It returns the current context (Wi-Fi networks) in JSON (JSONP) format.

Why do we need a web server? It lets us stay in the web development domain only. There is a simple and clear instruction for web masters:

- add SpotEx script to your page

<script type = "text/javascript" src = http://localhost:8080/spotex.js> </script>

- describe your info snippets as div blocks:

<div rel="spotex" net="WiFi_SSID" levelMin="" levelMax="">
Your HTML code
</div>

Our "old" rules could be presented via collection of attributes for HTML tags.

In this case, JavaScript code loaded from local server will be able to proceed all the div blocks related to SpotEx, and set visibility attributes depending on the context.

Such simple trick let us convert any existing HTML page into "Wi-Fi context aware" version. Note that if our script is not available, the page will work as a "standard" HTML page.

There is also a "side" effect for SpotEx application – WiFiChat service [19]. This mobile application uses the principles described in this article and offers communication tools (web chat and discussions groups) for mobile users nearby the same Wi-Fi access point. Think about it as "SpotEx with predefined content". The typical use case – we have Wi-Fi network in the train and this application automatically provides the discussions forum for the passengers. Or, keeping in mind that the "base" Wi-Fi network for this service could be opened right on the phone, this application can present personal forum (classified for example) as well as web chat for phone owner. This Android application is actually a wrapper for web mashup that combines HTML5 web chat engine and cloud based forums from Disqus:

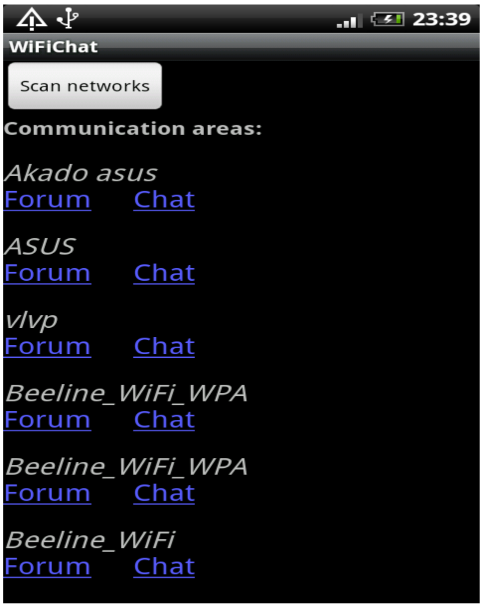


Figure 3. Wi-Fi chat application

It is the typical tool for the communications on the go. It is used right now in the public transport where some WiFi service is present.

IV. THE FUTURE DEVELOPMENT

Here, we see several almost obvious steps. At the first hand it is open API. In the current implementation SpotEx front-end actually obtains data in JSON (JSONP) format from our server-side database.

As soon as API is going live, the next step is almost mandatory. It should be the stuff that will simplify the development. The good candidates here are web intents [20] Web Intents is a framework for client-side service discovery and inter-application communication. Services register their intention to be able to handle an action on the user's behalf. Applications request to start an action of a certain verb (for example share, edit, view, pick, etc.) and the system will find the appropriate services for the user to use based on the user's preference. It is the basic.

Intents play the very important role in Android Architecture. Three of the four basic OS component types - activities, services, and broadcast receivers - are activated by an asynchronous message called as intent.

Intents bind individual components to each other at runtime (you can think of them as the messengers that request an action from other components), whether the component belongs to your application or another.

Created intent defines a message to activate either a specific component or a specific type of component - an intent can be either explicit or implicit, respectively.

For activities and services, an intent defines the action to perform (for example, to "view" or "send" something) and may specify the URI of the data to act on (among other things that the component being started might need to know). For example, our intent might convey a request for an activity to show an image or to open a web page. In some cases, you can start an activity to receive a result, in which case, the activity also returns the result in an Intent (for example, we can issue an intent to let the user pick a list of nearby images and have it returned to us - the return intent includes data in some format)

Going to our context aware browsing it means that our mobile devices will be able to present local data without low-level programming.

Web Intents puts the user in control of service integrations and makes the developers life simple.

Here, is the modified example for web intents integration for the hypothetical web intents example:

1. Register some intent upon loading our HTML document

```
document.addEventListener("DOMContentLoaded",
function() {
        var regBtn = document.getElementById("register");
        regBtn.addEventListener("click", function() {
window.navigator.register(http://webintents.org/spotex,
undefined);
        }, false);
```

2. Start intent's activity and pass it extra data (context info)

```
        var startButton =
document.getElementById("startActivity");
        startButton.addEventListener("click", function() {
          var intent = new Intent();
          intent.action = "http://webintents.org/spotex";
          intent.putExtra("WiFi_List", List_Of_Networks);
            window.navigator.startActivity(intent);
        }, false);
```

3. Get local info snippets (note – in JSON rather than XML) and display them in our application

```
        window.navigator.onActivity = function(data) {
          var output = document.getElementById("output");
          output.textContent = JSON.stringify(data);
        };
      }, false);
```

Obviously, that it is much shorter than the long sequence of individual calls as per any Open API.

The key point here is *onActivity* callback that returns JSON formatted data. Additionally, web intents based approach is asynchronous by its nature, so, we don need to organize asynchronous calls by our own.

Also, we are planning to add Bluetooth measurements too. But, by our vision, we should avoid the typical Bluetooth usage cases and do not use push proxy as per classical Bluetooth marketing. We think that the end users do not welcome at least push approach and it is the source of problems with Bluetooth proximity. Vice versa, in SpotEx Bluetooth nodes will be used the same manner we are using Wi-Fi access points – as presence triggers. In other words, we will add the ability to describe rules for Bluetooth nodes too.

SpotEx approach could be extended also towards accumulating some ideas from the collaborative locations. We can add trilateration terms (conditions) to our rules, but present them in terms of fuzzy logic (close than, relatively close, etc.) It helps us incorporate grid data in case of many devices without any infrastructure preparation.

The next area we are going to pay attention to is Wi-Fi Direct specification. Wi-Fi Direct devices can connect directly to one another without access to a traditional network, so mobile phones, cameras, printers, PCs, and gaming devices can connect to each other directly to transfer content and share applications anytime and anywhere. Devices can make a one-to-one connection, or a group of several devices can connect simultaneously. They can connect for a single exchange, or they can retain the memory of the connection and link together each time they are in proximity [21].

As per Wi-Fi Direct spec, a single Wi-Fi Direct device could be in charge of the Group, including controlling which devices are allowed to join and when the Group is started. All Wi-Fi Direct devices must be capable of being in charge of a Group, and must be able to negotiate which device adopts this role when forming a Group with another Wi-Fi Direct device. The device that forms the Group will provide the above described dynamically assembled web page with discovered services. It is how SpotEx could be extended to Wi-Fi Direct.

V. CONCLUSION

This paper described a new context-aware browsing model for mobile users based on the ideas of Wi-Fi proximity. Developed service can use any existing as well as the especially created (described) Wi-Fi network as presence trigger for discovering user-defined content right to mobile subscribers.

Proposed approach is completely software based and does not require additional hardware investments. For using SpotEx you need nothing except the smart phone. Also this approach supports ad-hoc solutions and despite existing Wi-Fi based indoor-location techniques does not require the upfront space preparations.

This service could be used for delivering commercial information (deals, discounts, coupons) in malls, hyper-local news data, data discovery in Smart City projects, personal news sites, etc.

REFERENCES

[1] C. Seeger, A. Buchmann, and K. Van Laerhoven: Wireless sensor networks in the wild: three practical issues after a middleware deployment Proceedings of the 6th International Workshop on Middleware Tools, Services and Run-time Support for Networked Embedded Systems ACM New York, NY, USA ©2011 ISBN: 978-1-4503-1069-7

[2] R. Meier, V. Cahill, A. Nedos, and S. Clarke: Proximity-Based Service Discovery in Mobile Ad Hoc Networks Distributed Applications and Interoperable Systems Lecture Notes in Computer Science, 2005, Volume 3543/2005, pp. 1147-1149, DOI: 10.1007/11498094_11

[3] M. Weiser: "Ubiquitous Computing," IEEE Hot Topics, vol. 26, pp. 71-72, 1993.

[4] S. Poslad Ubiquitous Computing: Smart Devices, Environments and Interactions DOI: 10.1002/9780470779446.fmatter Copyright © 2009 John Wiley & Sons, Ltd

[5] C. Bolchini1, G. Orsi, E. Quintarelli, F. A. Schreiber, and Letizia Tanca: Context modeling and context awareness: steps forward in the context-addict project. IEEE Data Eng. Bull., 34(2): pp. 47–54

[6] R. Ahas, S. Silm, O. Järv, E. Saluveer, and Margus Tiru: Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones; Journal of Urban Technology Volume 17, Issue 1, 2010 pp. 3-27

[7] Comparison of Wireless Indoor Positioning Technologies http://www.productivet.com/docs-2/Wireless_Comparison.pdf Retrieved: Feb, 2012

[8] F. Lassabe, P. Canalda, P. Chatonnay, and F. Spies “Indoor Wi-Fi positioning: techniques and systems” Annals of Telecommunications Volume 64, Numbers 9-10, pp. 651-664, DOI= 10.1007/s12243-009-0122-1

[9] C. Fretzagias and M. Papadopouli, Cooperative Location-Sensing for Wireless Networks, Proceedings of the Second IEEE International Conference on Pervasive Computing and Communications (PerCom'04), p.121, March 14-17, 2004

[10] R. José, A. Moreira, H. Rodrigues, and N. Davies: The AROUND Architecture for Dynamic Location-Based Services in Mobile Networks and Applications Volume 8, Number 4, pp. 377-387, DOI: 10.1023/A:1024531629631

[11] AllJoyn https://www.alljoyn.org/ Retrieved: Mar, 2012

[12] Nokia Instant Community http://research.nokia.com/news/9391 Retrived: Feb, 2012

[13] LocalSocial http://www.rococosoft.com/products/localsocial Retrieved: Mar, 2012

[14] SpotEx http://spotex.linkstore.ru Retrived: Feb, 2012

[15] D. Namiot and M. Schneps-Schneppe: About location-aware mobile messages International Conference and Exhibition on. Next Generation Mobile Applications, Services and Technologies (NGMAST), 2011 5th International Conference on Issue Date : 14-16 Sept. 2011 pp.: 48 – 53 ISSN : 2161-2889 Print ISBN: 978-1-4577-1080-3 DOI : 10.1109/NGMAST.2011.19

[16] E. Friedman-Hill Jess in Action: Java Rule-based Systems, Manning Publications Company, June 2003, ISBN 193011089

[17] C. L. Forgy "RETE: A fast algorithm for the many pattern/many object pattern match problem", Artificial Intelligence 19(1): pp. 17-37, September 1982

[18] P. J. Brown and G. J. F. Jones.:Context-aware retrieval: Exploring a new environment for information retrieval and information filtering. Personal Ubiquitous Computing, 5(4): pp. 253–263, 2001.

[19] Wi-Fi chat http://wifichat.linkstore.ru Retrieved: Mar, 2012

[20] Web Intents: http://webintents.org/ Retrived: Feb, 2012

[21] Wi-Fi CERTIFIED Wi-Fi Direct http://www.cnetworksolution.com/uploads/wp_Wi-Fi_Direct_20101025_Industry.pdf Retrieved: Feb, 2012

Mental Health Engagement Network (MHEN)

Connecting Clients With Their Health Team

Dr. Cheryl Forchuk, Associate Director
Nursing Research
Lawson Health Research Institute
London, Canada

Dr. Abraham Rudnick, Associate Professor
Departments of Psychiatry and Philosophy
University of Western Ontario
London, Ontario

Dr. Jeffrey Hoch, Health Economist and
Research Scientist,
St. Michael's Hospital
Toronto, Canada

Mike Godin, Team Leader
Housing Program
Canadian Mental Health Association
London, Canada

Dr. Lori Donelle, Assistant Professor
Faulty of Health Sciences
University of Western Ontario
London, Canada

Dr. Diane Neal, Assistant Professor
Faculty of Information & Media Studies
University of Western Ontario
London, Canada

Dr. Robbie Campbell, Associate Professor
Department of Psychiatry and Physician Lead
Regional Mental Health Care
London, Canada

Walter Osaka, Peer Specialist
CanVoice
London, Canada

Betty Edwards, Peer Specialist
CanVoice
London, Canada

Dr. Elizabeth Osuch, Associate Professor
University of Western Ontario
London, Canada

Dr. Ross Norman, Professor
Department of Psychiatry
University of Western Ontario
London, Canada

Dr. Evelyn Vingillis, Professor
Schulich School of Medicine and Dentistry
University of Western Ontario
London, Canada

Dr. Beth Mitchell, Director
Mental Health Program
London Health Sciences Program
London, Canada

Dr. Jeffrey Reiss, Professor & Chair
Department of Psychiatry, University of Western Ontario,
Site Chief, Mental Health Program, LHSC
London, Canada

Mike Petrenko, Executive Director
Canadian Mental Health Association
London, Canada

Dr. Deb Corring, Administrative Psychiatric Lead
Mental Health Transformation
Regional Mental Health Care
London, Canada

Abstract— **This research study introduces, delivers, and evaluates the benefits of using web and mobile technology to provide consistent supportive health care to individuals living within the community who have been diagnosed with a mental illness. This longitudinal, mixed method study will consist of 400 individuals who have been diagnosed with either a mood or a psychotic disorder who are currently working with mental health care professionals in the community. The participants will have access to a personalized health record and applications related to their care plan on a smart phone. It is hypothesized that the use of smart technologies in the treatment of mood and psychotic disorders will improve quality of life while reducing health care costs through a reduction in hospitalizations and hospital room visits. Baseline findings will be available in May 2012.**

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

Keywords- Smart technology; Mental Health Care; Personal Health Records; Quality of Life; Health Care Costs

I. INTRODUCTION

The economic cost of mental illness to Canada was recently estimated at $51 billion annually, and still the current "system" of care is fragmented and without sufficient resources (financial, human, and technological), per "The Healthcare Interview" Canadian Healthcare Network, December 2009 [1]. The 2009 estimations showed a dramatic increase from 2003, when it was estimated that the Canadian economic burden due to mental illness was ~ $34 billion ($1,056 per capita), which was a 3-fold increase from 1998 estimations of $12.3 billion. Twenty percent of Canadians will personally experience a mental illness in their lifetime and most others will experience mental illness indirectly through a family member, friend, or colleague. Mental illness affects people of all ages, regardless of education, income level, or culture [2]. With mental health services being at a crisis point, even though billions of dollars are being invested to help, most people in need of care will not receive the care they need [3].

Commonly, mental health care is prioritized for individuals with the most severe symptoms, due to the lack of available services. This runs counter to the general health care system where prevention and intervention at less serious levels is the norm. The current way of treating psychiatric illness is unsustainable and only through developing new service models that provide support and early intervention, will the mental health care system be sustainable. A new and potentially more sustainable method of providing mental health care would be that of employing smart technologies to enhance the treatment of mental health clients.

Several applications, both mobile and web based, have been developed employing smart technology to support health care. Examples include: My Mood Monitor [4] and MedHelp's Mood Tracker [5], which are both applications designed to monitor a user's mood; medication adherence assistance applications such as RxReminder [6]; the Mobile Assessment and Treatment for Schizophrenia (MATS) [7] which is an interactive text-messaging intervention; and the use of Personal Health Records/Electronic Health Records. However, the effectiveness of these applications has not been extensively researched, particularly within mental health care, nor have these applications ever been utilized in conjunction with each other to provide a holistic care package. This study will be the first of its kind to employ all of these applications in one complete platform while also including the use of personalized care plans which have previously not been utilized in this fashion.

The Mental Health Engagement Network (MHEN) is focused on putting technology in the hands of clients of the mental health system and their clinical team to demonstrate how to more effectively and efficiently deliver health care services. This project will deploy TELUS health space™ consumer health platform along with a customized personal health record application and interactive tools that support a novel way to provide patients with standardized health services, ongoing monitoring and regular communication with their care team. This innovative solution will help coordinate care across the continuum; ensuring that services are more accessible, patient-centered, and promote the empowerment of individuals so they can better manage their own health. From a population perspective, this proposed system re-design will have the capability to reduce or prevent acute episodes of mental illness and reduce the severe pressures on an already over burden health care system.

This document proceeds as follows. Section II describes the methods employed by the MHEN project. The expected outcomes are outlined in Section III and the conclusion is presented in Section IV.

II. METHODS

The following subsections describe the methods of theMHEN project .

A. Study Design

This delayed implementation research study launched in September 2011 and will conclude in November 2013. It includes 50 community mental health providers and 400 research community clients who have been diagnosed with a mood disorder or a psychotic disorder. Clients will be randomized into Group 1 (early intervention) or Group 2 (later intervention). Group 1 (200 participants) will receive a handheld device, a TELUS health space™ account, and version 1.0 of the Lawson SMART record during Phase I. The remaining 200 clients, Group 2, will initially act as a control group, and at Phase II (6 months later) will receive a handheld device, a TELUS health space™ account, and version 2.0 of the Lawson SMART record. Version 2.0 of the Lawson SMART record will be based on the feedback acquired from the care providers and clients in Group 1 who have participated in focus groups. Subjects will be recruited from local community programs and will be randomly assigned to the early or delayed adoption groups.

The specific electronic tools proposed in this project will leverage the use of web and mobile devices to provide cues to support daily structure to individuals. The tools will include self-assessment applications such as mood and medication monitoring. These tools will be linked in real-time to action plans designed by the consumer to be used when assistance is required in relation to their mental health. It is proposed that the action plans be linked with an individual's crisis plan, so prompts (i.e. SMS texts sent directly to the individual's mobile device) can be sent automatically to the individual based on their recovery plan activities. Through the use of communication tools, self-care will be encouraged and supported, allowing for a sense of empowerment in the consumer. These activities will be managed through the individual's personalized care plan, deployed through TELUS health space™, along with a Lawson SMART record to empower the consumer with their own health care information. This application will provide the consumer with the ability to share their own health care data, securely with their multiple care providers. Care providers will then have the ability to provide treatment based on prior care, so the consumer receives continuity of care. This will allow care providers to avoid costly duplications of tests and interventions that have already been determined unsuitable.

In summary, the present project will introduce, deliver, and evaluate the benefits of empowering consumers with their personal health information, while using leading-edge communication technologies to deliver a more consistent type of supportive care to individuals who most urgently need it.

B. Sample

The 400 client subjects and 50 care provider subjects will be recruited through community programs at London Health Sciences Centre and St. Joseph's Health Care, London as well as through the Canadian Mental Health Association (London-Middlesex Branch) and WOTCH Community Mental Health Services.

C. Data Collection

Individual interviews with client participants will be held during the baseline period and then every 6 months for 18 months. Data collected will include use of mobile devices, perception of usefulness, quality of life, general health, and use of health and social services (including hospitalizations and emergency room visits as well as other services). The knowledge learned through Group 1, over the initial 6 months, will provide baseline and comparative data to understand the client's perceptions for designs and outcome purposes of the technology intervention. Improvements in the technology approaches made during the initial 6 month period will enhance the tools for Group 2. We are expecting different feedback from groups, one to guide development and one to improve what is developed.

Focus groups will take place at multiple points during the study. Group 1 will participate in three focus groups: (1) approximately one month after receiving the handheld device and a TELUS Health Space account to discuss usability and adoption; (2) follow up focus groups will be held two months thereafter to discuss the benefits and pitfalls associated with the technology and to form base recommendations for the next phase of study (Group 2); and (3) 6 months later to discuss future recommendations. Group 2 will also participate in 3 focus groups. They will meet in a similar timeframe: (1) approximately one month after receiving the handheld device and a TELUS Health Space account to discuss usability and adoption; (2) follow up focus groups will be held two months thereafter to discuss the benefits and pitfalls associated with the technology; and (3) 6 months later to discuss future recommendations. Focus groups with staff/health care providers will also be held so that issues can be identified and addressed quickly.

III. PRELIMINARY RESULTS

Preliminary quantitative data analysis of 123 client subjects (61 men and 62 women) indicates that the mean age of the population is 40.88 (SD = 12.814). The most common psychiatric diagnosis in this sample population are mood disorders (69.1%), psychotic disorders (54.5%), anxiety disorders (41.5%), substance-related disorders (10.6%), personality disorders (8.9%), disorders of childhood/adolescence (7.3%) and developmental handicaps (0.8%). Most client subjects indicated that they have been admitted to the psychiatric hospital at least once (87%) and of those individuals, most have been admitted a mean of 8.83 times (SD = 13.772).

Initial quantitative analysis shows that most client subjects are generally comfortable with the use of technology. Most indicated that they were either extremely comfortable (22.0%) or slightly comfortable (25.2%) with technology generally, while only a minority said that they were either slightly uncomfortable (4.9%) or extremely uncomfortable (5.7%) with technology generally. Despite this general level of comfort, only a minority of participants had regular access to a computer at home (38.2%) and only a slightly larger population (41.5%) owned a cell phone. These findings suggest that client subjects will be receptive to training on the use of smart technologies, but also that they often do not have regular access to smart technologies.

IV. EXPECTED RESULTS

Baseline data will be completed by April 2012 and the training and implementation of the first 200 participants will be conducted in May 2012.. For the conference we will be able to describe the sample and their initial perceptions of technology as well as describing the elements of the programs and training.

The overall hypothesis is smart health information technology (HIT) will improve quality of life and reduce health care system costs. To test this hypothesis we will use an evaluation framework that includes four levels of analysis: effectiveness, economic, ethical and policy. Development and testing of a more cost-effective means of addressing mental health issues will increase the ability to provide the best practice at an affordable cost which benefits consumers and taxpayers.

The current study will be done in a geographical setting that represents a microcosm of environments found across Canada, making research done within this region easily translational throughout Canada. It includes a large integrated mental health focus to evaluate smart technology within the continuum of care from early intervention (Primary Care; Emergency and Acute Care; Community Integration; Tertiary Care to Rehabilitation). The study will provide an enterprise system solution to support health technology innovations, using mobile technologies and tools that will provide personalized, preventative consumer health care. The system re-design includes privacy security solutions developed by academic health researchers embedding knowledge into novel community based health tools, making tools directly translational for use in health care settings. This will also give researchers and health care providers an understanding of mental health client acceptability to receive personalized care through smart devices which has not been previously explored.

V. CONCLUSION

Preliminary results, including baseline data and initial focus groups, will be available in May 2012. It is expected that the use of smart technologies in the treatment of mood and psychotic disorders will improve quality of life while reducing health care costs through a reduction in hospitalizations and hospital room visits.

REFERENCES

[1] McAllister, J. (2009). The healthcare interview: Micheal Kirby breaks down barriers to better mental health care. *Canadian Health Care Network*.

[2] Canadian Mental Health Association. (2012). *Understanding mental illness.* Retrieved December 16, 2011, from http://www.cmha.ca/bins/content_page.asp ?cid =3&lang=1

[3] StatsCanada.(2004).*Canadian community health survey mental health and well-being.* Retrieved December 15, 2011, from http://www.statcan.gc.ca/pub/82-617-x/index.html

[4] M3 Information. (December 20, 2011.). *My mood monitor – Monitor my mood.* Retrieved from http://www.mymoodmonitor.com/

[5] MedHelp.(December 21, 2011). *Track your mood: Get your own personal mood tracker.*Retrieved from http://www.medhelp.org/land/mood-tracker

[6] Sterns, A. and Collins, S. (2007). Transforming the personal digital assistant into a useful health-enhancing technology for adults and older adults. *Generations, 28*(4).

[7] Granholm, E., Ben-Zeev, D., Link, P., Bradshaw, K., and Holden, J. (2011). Mobile assessment and treatment for schizophrenia (MATS). *Schizophrenia Bulletin*.

Experimenting with Instant Services Using NFC Technology

Anders Andersen and Randi Karlsen
Department of Computer Science
Faculty of Science and Technology, University of Tromsø
9037 Tromsø, Norway
Email: {Anders.Andersen,Randi.Karlsen}@uit.no

***Abstract*—In the NFC City project Near Field Communication (NFC) technology is used to demonstrate instant services. NFC makes it possible to connect electronic devices and initiate services with simply touching a tag (or a device) with a phone. In the NFC City project this is utilized in several experimental applications. A presenter application use this to set up a presentation (on a projector or screen) without having to connect cables or transfer files manually using USB memory sticks or similar. A chat application sets up a chat simply but touching two NFC enabled mobile phones. A party application transfers images and modifies the music play list by touching an NFC tag with the phone. A guide application provides information and suggest next sightseeing spot in the same way. The usage of NFC in these and a few other applications is presented and discussed, and shows how NFC can be used to provide instant services.**

***Keywords*-Mobile computing, Pervasive computing, Context-aware services, RFID tags**

I. Introduction

NFC (Near Field Communication) is a short-range wireless technology that tries to harmonize today's diverse contactless technologies, enabling current and future solutions like access control, ticketing, payment, loyalty programs, discount coupons, information collection and exchange. NFC is an extension of Radio Frequency Identification (RFID) technology. One way it differs from RFID is that it limits the range of communication to within 4 centimeters.

The use of NFC makes it possible to connect electronic devices and initiate services with a simple touch. It is believed that NFC will make people's lives easier and more convenient by enabling more intuitive access to new media and content services. It will for example be easier to pay for things; easier to discover, synchronize and share information; and easier to use transport and other public services.

This paper reports on the development and testing of a number of NFC-based applications [1]. In a previous short papert and poster we focused on social networking and context sensitivity [2]. In this paper we examine how various services can be made instantly available, in a simple manner, through the use of NFC-technology. We evaluate both the experience of implementing the NFC-based applications, and the experienced simplicity of using these applications.

The NFC City project is supported by the Norwegian research council, VERDIKT project number 201377.

Our applications vary widely, from "NfcPresenter", which uses NFC to simplify the process of starting presentations from your mobile device, to "NfcSafari", which uses the user's location to identify the closest sightseeing spot, and from there on takes him on a city safari. The applications "PartyShare", "Game Collector" and "Are You the One?" explores how powerful NFC can be in social settings. All applications, including "BluetoothChat", demonstrates how instant service can be achieved with NFC.

This work is part of an NFC City project, where the objective is to promote the development and use of NFC-based services. This objective will be achieved through the establishment of an NFC ecosystem including network infrastructure, trusted service manager (TSM) functionality, NFC services and user handsets. The project is also developing a tool set for the development of NFC services.

In the next two sections, related work and NFC technology is discussed. In Section IV, the tool set and the application experiments will be presented. In Section V, the presented work is discussed and evaluated, and finally, in Section VI, the paper is summarized and future work is discussed.

II. Related work

Many related projects developing and disseminating NFC-based applications exists. A multitude of NFC activities are reported in the NFC Forum [3]. Field trials with NFC based systems have been conducted in a number of cities over the world [4], including Nice (France), Madrid (Spain) and London (UK), and are spreading to more cities. In these cities different types of NFC applications are tested. Examples include ticketing, payment and information delivery applications.

NFC Forum has defined three different NFC operation modes; Card Emulation, Reader/Writer and Peer-to-Peer. In card emulation mode, the NFC device acts as (and eliminates the needs for) a physical object, such as a credit card, key, ticket or coupon. In reader/writer mode, NFC devices can read and write data from/to NFC tags, while in peer-to-peer mode, data can be transferred between two NFC devices. In [5], a number of applications have been examined and classified according to the three operation modes.

The card emulation applications include payment [6], ticketing [7], mobile coupons [8], and loyalty applications [9]. Some

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

typical reader/writer applications are designed for delivering information about tourist attractions [10], [11], providing users with location-based information and helping them finding points of interest. In [12], NFC technology supports personalized advertising that combines user profiles on mobil devices and direct interaction displays, and in [13] NFC is used for exchanging pictures between mobile device and computer.

Applications supporting social networking are described in [14], [15]. In [14], NFC technology is used to ease the update of presence information on social networks, while in [15] a peer-to-peer mode application allows users to make friends by exchanging personal information between devices. The peer-to-peer mode is also used in [16], where NFC is used to build ubiquitous games.

III. NFC TECHNOLOGY

NFC-Forum defines the technology as follows: *Near Field Communication (NFC) technology makes life easier and more convenient for consumers around the world by making it simpler to make transactions, exchange digital content, and connect electronic devices with a touch.*

NFC can be implemented in all types of devices, including your smart phone, your tablet, your computer and your car. However, the number of NFC-enabled devices today are quite limited. There has been a deadlock situation where handset vendors, service providers and users are waiting for the others to make a move. The lack of NFC handsets has hindered service development and lack of services has not been good incitement for developing new handsets. However, in current releases of mobile handset models we see a significant increase in NFC enabled devices.

In short, NFC is a set of short-range wireless technologies. Where Bluetooth and RFID may work within a range of up to 100 meters, NFC is limited to only work within distances as short as 4 cm. Also, where Bluetooth has a maximum bandwidth of 2.1 Mbit/s, NFC uses a bandwidth ranging from 106 kbit/s to 424 kbit/s. NFC operates at 13.56 MHz on the ISO/IEC 18000-3 air interface.

NFC always involves an initiator and a target; the initiator actively generates an RF field that can power a passive target. This enables NFC targets to take very simple form factors such as tags, stickers, key fobs, or cards that do not require batteries. NFC peer-to-peer communication, where both devices are powered, is also possible. However, the limited bandwidth makes it inferior to Bluetooth for big data transfers.

NFC tags contain data, and may be configured to be both read-only and rewritable. They can be custom-encoded by their manufacturers or use the specifications provided by the NFC Forum, including the Type 2 Tag format [17] for the tag header and the NDEF data format [18] for the payload. The tags can securely store personal data [19] such as debit and credit card information, loyalty program data, PINs and networking contacts, among other information. The NFC Forum defines four types of tags, which provide different communication speeds and capabilities in terms of configurability, memory, security, data retention and write endurance.

The NFC Data Exchange Format (NDEF) specification [18] defines a message encapsulation format to exchange information between NFC Forum Devices and NFC Forum Tags. In the specification it is described as follows: *NDEF is a lightweight, binary message format that can be used to encapsulate one or more application-defined payloads of arbitrary type and size into a single message construct. Each payload is described by a type, a length, and an optional identifier.* The type identifiers may be URIs, MIME media types, or NFC-specific types.

IV. THE EXPERIMENTS

All Android applications have been developed using Android SDK Tools, Revision 11, which includes the Android SDK and the Android Development Tools (ADT). NFC functionality is implemented using the Android API package *android.nfc*. All applications developed on the Android platform have been tested on a Samsung Nexus S with Android version 2.3.4.

The NFC cards and stickers used are 64 bytes MiFare Ultralight tags. These fulfill the NFC Forum type 2 tag specification with a total payload size of 48 bytes. The desktop NFC development has been on a SCM Microsystems SCL3711. This is a USB device with a PN532 chip, which supports emulation of all NFC Forum type tags and many other non-standard tags. It also supports read/write communication with the most common tag types. We have used Libnfc [20] development branch Revision 1120 for this development. Libnfc is an open source SDK and programmers API for working with NFC on traditional desktop computers.

A. EasyNfc Tool Set

As Libnfc is a low level API, our focus has been on developing higher level APIs for use in future application development processes. We have developed EasyNfc, which is a library that simplifies the process of developing NFC applications on the Android platform. The library contains extendible Activity classes for reading and writing to tags, P2P, and Bluetooth pairing of Android devices. It elegantly reduces the code required when creating NFC applications by hiding much of the boiler plate code from the application developer. This is done by creating EasyNfc as an extension of the original Activity-class. The application developer can simply extend one of the EasyNfc classes to gain access to a powerful set of NFC tools.

Our APIs include functionality for emulating and reading Type 2 Mifare Ultralight tags. It also simplifies the process of creating and parsing NDEF messages. The implementation of these APIs are done in C, but we have implemented a Python wrapper to make it even easier to integrate with our services in future projects. The advantage of emulating an NFC tag with a USB device compared to using a static tag is the possibility for varying responses. As the tag is emulated in software the application can give different responses dependent on context. The tag data may vary over several parameters, like what time of the day it is or what user is reading it. This may prove

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

useful in many areas, including context sensitive functionality, permission control and security.

We have also included an Android application, called "Tagger", in the tool set. Tagger is a standard NFC tag reader and writer. The application utilizes the NFC capabilities of the mobile device to read Type 2 NDEF tags and display them in plain text to the user. The reader is capable of reading tags with Plain Text [21] and URI [22] payload records. Tagger can also write to tags, but this is limited to Plain Text records. As Tagger is implemented complying to the Type 2 and NDEF standards, the tags written with the application can be read by all standard compliant readers. The work done on Tagger provides the foundation for the other Android applications that have been build.

B. NfcPresenter

NfcPresenter is an Android application that simplifies the user-computer interaction by creating a more intuitive way to present slide-show presentations. Traditionally, starting a presentation is a cumbersome process. The person doing the presentation usually has two alternatives; either connecting his own laptop to the projector, or transfer the presentation with a USB memory stick to a local computer, which is already connected. Both routines are unnecessary tedious.

In the first case, inability to establish connection between laptop and projector and not being able to display the presentation correctly, are two common and often time-consuming problems. The alternative, transferring the presentation to a USB memory stick, connecting the USB device to the local computer, navigating to the relevant file, opening it, and starting the presentation, is not any more tempting.

With NfcPresenter we aim to create a smooth and painless presentation experience by utilizing NFC technology. Starting a presentation can be done by selecting the appropriate presentation file on the mobile device and touching it to an NFC tag present at the presentation location. Simple as that. When the presentation has started, the Android device serves as a remote control. With a swipe you can show the next or previous slide. In addition each slide is shown on the screen of the mobile device for better control.

NfcPresenter also lets a user download the currently active presentation to a mobile device. In other words, a spectator of the presentation can use the application to download the presentation to his own device by scanning a tag. This is a useful feature for distributing presentation material.

As explained, NfcPresenter utilizes NFC to collect the URL from a tag present at presentation location. The URL directs to a server running on the local presentation computer, which is already connected to the projector. The application uses the recently acquired URL to upload the presentation file to the server. The server starts the presentation and can now answer to other commands to control it remotely.

We predict that the way presenting is handled in NfcPresenter is the future for a presentation setup. To include NfcPresenter-like server functionality into popular end-user products like Microsoft Powerpoint or Apple Keynote would be a simple and efficient way to distribute this kind of features. It could be a great enhancement to the current presenting experience.

(a) NfcPresenter (b) PartyShare

Fig. 1. NFC-based applications

C. BluetoothChat

BluetoothChat demonstrates instant setup with NFC. It is a simple chat client where the user use the NFC features of their mobile phone to set up a bluetooth chat between the handsets. No other configuration is needed. The users open the chat program and then touch the other user's phone. NFC is then used to exchange the information needed to set up the bluetooth communication between the two phones. The communication is the set up without any user involvement, and the chat can start immediately. The advantage of this approach is not only a simplified setup, but also that the two users can be sure that they are actually communicating with the right person. The physical presence (at setup) and the short range of NFC ensure this.

This application can be easily extended to set up chat or other forms of communication using other network technology than the relative short range bluetooth network. In such a context the setup can be done once (when the users meet physically) and used later again and again remotely.

D. PartyShare

The goal of PartyShare is to bring multimedia from handheld devices to stationary devices in a seamless manner. It consists of a client running on an NFC enabled Android device, and a server running on a computer. The application allows you to share music or images from the Android device to the computer in real-time. The transfer is done simply by touching the Android device to the stationary device. As a result, the images are viewed and the music is played immediately.

In principle, PartyShare uses the HTTP protocol to transfer the multimedia from the Android device to the computer. This

Copyright (c) IARIA, 2012.

implies that the application could have been built without NFC technology. However, configuring such an application would be tedious. For instance, moving from one location to another would require the user to enter a new IP-address in order to interact with the current stationary device.

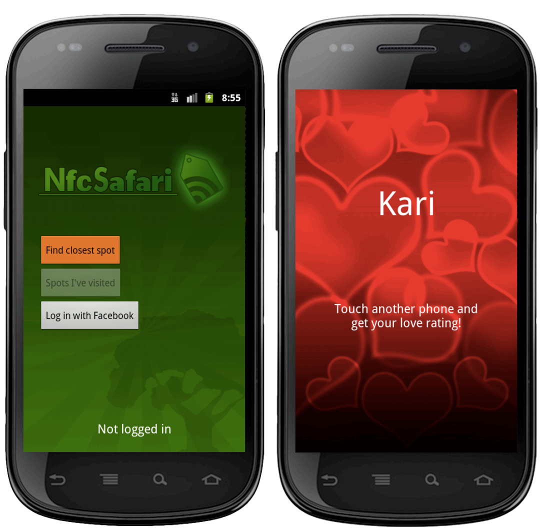
Entering a URL on a handheld device is both time consuming and cumbersome. By taking advantage of NFC technology we can remove the configuration step completely, and as a result get a smart and intuitive sharing application. The trick is to store the relevant URL in the NFC tag. This way the Android application can read the required information directly from the relevant device, and get immediate feedback as his image is viewed, or the track is played.

The server is implemented in Python, using a share of open source python modules. The ImageViewer, which displays the latest image is created using Pygame, and the music sharing is implemented using a self made libspotify-wrapper (spotipy). The web server responds to HTTP post requests on the URIs `/upload/image` and `/upload/track`.

On the client side we use internal Android activities to handle the camera and the gallery to choose image, and the Spotify meta data API to choose music. The music sharing works by transferring a Spotify URI, which is then used by spotipy to play the song.

E. NfcSafari

NfcSafari is a tour guide application that aims to help people discover, experience and remember interesting locations and spots. The general application area is giving tourists visiting cities and locations for the first time a helping hand by guiding them to interesting sightseeing spots. The overall goal of creating the application is to explore how NFC can be utilized in pervasive applications.



(a) NfcSafari (b) Are You The One?

Fig. 2. NFC-based applications

The application initiates the tour by using the current location of the user to find nearby spots. The user is given a choice between the three closest spots, and the user will start the Safari by choosing one of them. The next screen shows a map, guiding the user from his current location to the spot he chose.

When the user finds a spot, he needs to touch the nearby NFC tag to register his presence. By reading the ID of the spot, the application knows what web service to contact to download a description of the spot. If the user is logged in, we can now choose to share information about the spot through Facebook. It is also possible to give the spot a rating from 1 to 5 stars. When the user is happy with the experience he can continue the tour. The next spot is shown on the map, and the process starts from the top. All visited spots are stored in a local database.

In NfcSafari all sightseeing spots are predefined and grouped on the server. The idea is to let spots in close vicinity to each other, or spots with the same theme, to be placed in the same tour. All tours are defined in XML files, where each file describes a single Safari.

Each Safari has an ID and a name, in addition to having a list of spots with names, IDs and location. When the server starts, it parses all the Safari XMLs for spots and locations, which in turn is used when the Android application requests the closest spot. When this happens, the server computes the distance from the user to all active spots (from all Safaris) and returns a list of the three closest of them. When the user chooses one of them, the Safari it belongs to is chosen as the active tour and the corresponding XML is downloaded. When the user is done with the current spot, the next spot in the current Safari is chosen. Communication between the server and the Android client is done over HTTP, and the server is implemented using web.py.

F. GameCollector

Game collector is a prototype of a concept that can increase social interaction between mobile gamers. The idea is to make games, that are generally not made for social interaction, social.

In the prototyped example, the setting is a Pokemon like game, where the goal is to collect many collectables and use them in you game. With NFC, these collectables can be shared among friends giving a social experience to an otherwise not social game.

G. Are You The One?

"Are You The One?" is a quirky "love tester" that calculates the compatibility between two persons. The goal of the application is to illustrate how NFC can be used in social scenarios. The application works by having the user enters his or her name, and then touch another handset with the same application. As a result their "love match", or compatibility, is visualized on screen as a percentage.

The compatibility is calculated deterministically using their names, and obviously has no basis in reality. Nevertheless, it is a fun example of a small application that delivers something new by incorporating NFC as a vital component.

V. DISCUSSION

NFC is a relatively new technology, and does not have the infrastructure of other older technologies. Because of this, there is less support available for development of NFC services and applications.

Developing NFC application on Android is not complex, but current Android NFC implementations can be unstable during testing. This is both due to the immature NFC support in Android and the existing tool set. This is continuously improving and a thorough implementation and documentation of both new and legacy NFC standards has been made available.

Developing with Libnfc is challenging. Not only because its a low level API, but also because the documentation is a bit weak, and unforgiving to newcomers. To be able to use Libnfc, and to fully understand it, you will have to do a lot of research and understand a range of technical specifications. With few other alternatives, this is a challenge for NFC development on desktop/server computers.

The MiFare Ultralight tags used in this project had some limitations worth noting. The mobile device was unable to read the tags if they were too close to metal. This is definitely worth remembering when placing tag stickers.

A. Evaluation

NFC is useful in many areas, such as payment, information sharing [12], [23], security [19], [24], configuration, and social interaction [14]. It acts as a link between the physical and virtual world, giving us a simpler and more intuitive way of sharing information [25]. One of the properties of NFC is that context is implicit defined. This stems from the very short range of the tags. This can simplify many aspects of development, including security and configuration. Context wise NFC provides both location (location of tag), interest (selection of tag), time (when NFC device is used), and identity (owner/user of NFC device).

Apart from payment and ticketing, we believe that NFC for mobile devices are most useful where it can simplify the user-interactions of services or applications. The best example is setting up connections for WI-FI or Bluetooth, as exemplified in BluetoothChat. A task that can be tedious is replaced by an intuitive gesture that connects touching devices. We think that NFC is not introducing something entirely new in this context, but it can certainly create more convenience to already established applications and services. In some cases it can also improve the accuracy, for instance in the case of many points-of-interest at geographical location an NFC tag can be used to clarify where the current focus is.

A good example of this is NfcSafari, which does not accomplish something we could not have done using GPS more extensively. However, the GPS would limit us to outdoor spots and might not provide the necessary accuracy. QR codes are also an alternative to NFC in this setting, however, scanning QR codes is slightly less convenient as it takes more time and can be hard in the dark. If this application was to be launched in todays market it would have been wise to incorporate both QR codes and NFC. This is due to the still limited use of NFC enabled devices.

NfcPresenter underlines the convenience of NFC enabled applications. To configure the upload and download URL by a physical touch instead of entering it manually makes a mobile presentation application go from something a few people will use to something everyone will want to have.

BluetoothChat both removes the inconvenience of the setup process and ensures that the chat is with the person you have met physically (and have touched phones with). It is easy to see this used in many other applications where interaction between people is included. If the initial NFC connection is used to agree upon encryption keys it can be used to set up a secure channel between the devices.

PartyShare illustrates how NFC can enhance a social event by giving everyone the ability to participate. This type of crowd sourcing has been possible before NFC, but the threshold to join the fun can now be significantly lowered by making the act of sharing something come down to a simple touch.

Applications like “Are You The One?” illustrates how an apparently exhausted concept gets a fresh feel by incorporating NFC. The act of touching two devices to get a love match (instead of typing in the names on one device) give a greater connection between the two participants. The application illustrates the easy exchange of information between two mobile devices, and the concept obviously opens up for exchanging personal information already available on the devices.

VI. SUMMARY AND FUTURE WORK

These initial NFC experiments of the NFC City project have explored NFC technology both in design of new services and in the development of actual applications. Our experience with implementation and use of these applications indicate that NFC has a role to play in a wide range of future applications. In these experiments we have demonstrated NFC used to establish instant services.

In general, there is a lot more to explore regarding NFC technology. While we have investigated possible application and service areas, traditional NFC research areas, such as mobile payment, security and authentication, has been left untouched. This was done consciously, as other NFC City partners are currently exploring these fields to great extent. The result of this ongoing work will later be incorporated into our NFC activities.

While we have implemented a number of prototype applications utilizing NFC technology on mobile platforms, we have not developed any desktop prototypes using the USB devices. An advantage of the USB device is the possibility to create tags dependent on context. Although this is achievable by utilizing static tags combined with URLs and web access, USB may be a better option as it delivers the same context dependent experience without relying on a WI-FI or 3G/EDGE connection for the mobile device.

VII. ACKNOWLEDGMENTS

Øyvind Holmstad and Tor Kreutzer have provided the main contribution to the implementation of the NFC experiments.

This includes the Android applications and the server side application support. These experiments would have been impossible without support from our project partners in the NFC city project. Especially the support from project leader Sigmund Akselsen and project member Peter Dahle (both at Telenor) has been important. The support from the technical staff at the Department of Computer Science at University of Tromsø has also been important during these experiments.

References

[1] Ø. Holmstad and T. Kreutzer, "NFC prototype bonanza," Department of Computer Science, Faculty of Science and Technology, University of Tromsø, Norway, Tech. Rep., Jul. 2011.

[2] A. Andersen, Ø. Holmstad, R. Karlsen, and T. Kreutzer, "NFC city context sensitive and social networking experiments," in *PDT'11, Proceedings of the Workshop on Posters and Demos Track at Middleware 2011*. Lisbon, Portugal: ACM, Dec. 2011.

[3] "The near field communication forum," [retrieved: March, 2012]. [Online]. Available: http://www.nfc-forum.org/

[4] "The contactless & NFC city league," [retrieved: March, 2012]. [Online]. Available: http://contactlesscities.wordpress.com/

[5] K. Ok, V. Coskun, M. Aydin, and B. Ozdenizci, "Current benefits and future directions of NFC services," in *Education and Management Technology (ICEMT), 2010 International Conference on*, Nov. 2010, pp. 334–338.

[6] M. Pasquet, J. Reynaud, and C. Rosenberger, "Secure payment with NFC mobile phone in the SmartTouch project," in *International Symposium on Collaborative Technologies and Systems (CTS 2008)*, May 2008, pp. 121–126.

[7] S. Ghiron, S. Sposato, C. Medaglia, and A. Moroni, "NFC ticketing: A prototype and usability test of an NFC-based virtual ticketing application," in *NFC'09, First International Workshop on Near Field Communication*, Feb. 2009, pp. 45–50.

[8] S. Dominikus and M. Aigner, "mCoupons: An application for near field communication (NFC)," in *21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)*, vol. 2, May 2007, pp. 421–428.

[9] B. Ozdenizci, V. Coskun, M. Aydin, and O. Kerem, "NFC loyal: A beneficial model to promote loyalty on smart cards of mobile devices," in *International Conference for Internet Technology and Secured Transactions (ICITST)*, Nov. 2010, pp. 1–6.

[10] E. Siira, T. Tuikka, and V. Törmänen, "Location-based mobile wiki using NFC tag infrastructure," in *First International Workshop on Near Field Communication (NFC'09)*, Feb. 2009, pp. 56–60.

[11] F. Borrego-Jaraba, I. Luque Ruiz, and M. Á. Gómez-Nieto, "A NFC-based pervasive solution for city touristic surfing," *Personal and Ubiquitous Computing*, vol. 15, pp. 731–742, Oct. 2011.

[12] S. B. Paul Holleis, Gregor Broll, "Advertising with NFC," in *Workshop on Pervasive Advertising and Shopping, in conjunction with the 8th International Conference on Pervasive Computing (Pervasive 2010)*, Helsinki, Finland, May 2010.

[13] K. Seewoonauth, E. Rukzio, R. Hardy, and P. Holleis, "Two NFC interaction techniques for quickly exchanging pictures between a mobile phone and a computer," in *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI'09. New York, NY, USA: ACM, 2009, pp. 39:1–39:4.

[14] A. Fressancourt, C. Hérault, and E. Ptak, "NFCSocial: social networking in mobility through IMS and NFC," in *First International Workshop on Near Field Communication*, Feb. 2009, pp. 24–29.

[15] E. Siira and V. Törmänen, "The impact of NFC on multimodal social media application," in *Second International Workshop on Near Field Communication (NFC'10)*, Apr. 2010, pp. 51–56.

[16] P. Garrido, G. Miraz, I. Ruiz, and M. Gomez-Nieto, "Near field communication in the development of ubiquitous games," in *International Conference for Internet Technology and Secured Transactions (ICITST 2010)*, Nov. 2010, pp. 1–7.

[17] NFC Forum, "NFC type 2 tag operation specification," NFC Forum, Technical Specification T2TOP 1.1, May 2006, [retrieved: March, 2012]. [Online]. Available: http://www.nfc-forum.org/

[18] ——, "NFC data exchange format (NDEF)," NFC Forum, Technical Specification NDEF 1.0, Jul. 2006, [retrieved: March, 2012]. [Online]. Available: http://www.nfc-forum.org/

[19] E. Haselsteiner and K. Breitfuß, "Security in near field communication (NFC)," in *Workshop on RFID Security*, Malaga, Spain, Jul. 2006.

[20] "Public platform independent near field communication (NFC) library," [retrieved: March, 2012]. [Online]. Available: http://www.libnfc.org/

[21] NFC Forum, "NFC text record type definition," NFC Forum, Technical Specification RTD-Text 1.0, Jul. 2006, [retrieved: March, 2012]. [Online]. Available: http://www.nfc-forum.org/

[22] ——, "NFC URI record type definition," NFC Forum, Technical Specification RTD-URI 1.0, Jul. 2006, [retrieved: March, 2012]. [Online]. Available: http://www.nfc-forum.org/

[23] M. Isomursu, P. Isomursu, and M. Komulainen-Horneman, "Touch to access the mobile internet," in *OZCHI 2008 Proceedings*, Cairns, QLD, Australia., Dec. 2008, pp. 17–24.

[24] M. Hutter and R. Togl, "Touch'n trust: An NFC-enabled trusted platform module," *IARIA International Journal On Advances in Security*, vol. 4, no. 1 & 2, pp. 131–141, 2011.

[25] G. Broll, W. Reithmeier, P. Holleis, and M. Wagner, "Design and evaluation of techniques for mobile interaction with dynamic NFC-displays," in *TEI'11 Proceedings*, ACM. Funchal, Portugal: ACM, Jan. 2011, pp. 205–212.

Exploring the Relationship Between Smart City Policy and Implementation

Ellie Cosgrave
Industrial Doctorate Centre in Systems
University of Bristol
Bristol, UK
Ec5226@bris.ac.uk

Theo Tryfonas
Industrial Doctorate Centre in Systems
University of Bristol
Bristol, UK
Theo.tryfonas@bris.ac.uk

***Abstract*— The implementation of smart technologies gives rise to concerns about how imposed technologies create genuine value to a city. City leaders are facing increasing pressure to meet ambitious targets with limited resources. Much of the private sector heralds 'smart city' solutions as the way forward for meeting these targets, but city leaders often find it difficult to align the technology solutions with the intended policy outcomes. This paper investigates the core themes within the field of smart cities and future city policy, in order to derive an interpretive conceptual model of the relationships between them. We find that devising appropriate solutions should be framed not only by policy goals but a wider understanding of local challenges and opportunities, as well as an interpretation of public value.**

Keywords- Future cities; policy implementation

I. INTRODUCTION

The 'information age' has driven a significant shift in nearly all aspects of modern life. In the past decade, we have seen a fundamental change in the way we work (networking through social media, distance working etc.), how we shop (online, price comparison), interact with family and friends (skype, social media), and our expectations of government (FixMyStreet, opendata). It has also heralded a new era of activism and community unity. "In the Arab Spring, social media facilitated action in the Middle East and North Africa (MENA) region, providing a free and accessible method of organising and coordinating demonstrations"[1]. This was echoed in the London riots and the subsequent cleanup operation [2].

"The networked information environment has dramatically transformed the marketplace, creating new modes and opportunities for how we produce and consume information" [3]. Companies like Facebook, Amazon and Google have capitalized on this opportunity by using information to provide value to their customers. These companies utilize information as a core asset, and leverage it to create products and services that respond to user desires and expectations. The current 'information marketplace' in cities already creates value for citizens as highlighted in recent reports such as "Information Marketplaces: The New Economics of Cities" [4]. Innovative products and services create jobs and support citizens in navigating and using the city in effective, educational and enjoyable ways. However the true value has not yet been quantified or captured by city leaders. Governments are struggling to realize the opportunities offered by ubiquitous information, 'smart' technologies, social media, and anytime, anywhere access. They are unable to articulate the value of the market within their own city, let alone the 'value chain' [5] in term of inputs, outputs and outcomes for citizens. In an interview, Emer Coleman (Deputy Director Digital Engagement Government Digital Services at the Cabinet Office) explained, "this requires new leadership from the public sector. Data surfaces political decisions."

This paper investigates the core themes within the field of smart cities and future city policy, in order to derive an interpretive conceptual model of the relationships between them. Section two sets out a methodology for investigation, sections three to six investigate the core themes through analysis of second-hand evidence. Finally, we introduce our model in section seven and set a further research agenda in section eight.

II. METHODOLOGY

Investigating local government decision making around 'smart cities' requires an understanding beyond the business case analysis for a given solution or investment. While a rigorous financial justification is important for smart city investments, it is also useful to consider the wider context of decision-making; the motivations for change, the value systems, and the political paradigm within which both long and short term decisions are made. Furthermore, "a city is an enormously complex and open-ended system, with many intertwining force fields influencing its form simultaneously" [6]. An investigation into single technology-solutions is largely obsolete if the wider context and relationships are not understood. In this light, we have undertaken a multi-disciplinary literature review of the way in which decisions are made in the public sector. Taking a grounded approach we extracted core themes and explored the relationship and flows between them to create an interpretive conceptual model. We then explored, validated and developed this model through consultation with a series of 'smart city' experts.

We intend to use this model to compliment our research into how cities should invest in the 'information economy' and 'smart cities.' Setting this research in its wider context of public sector decision making helps to ensure that research findings are directed towards recommendations that are politically acceptable, actionable and that are a genuine contribution to the creation of public value.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

III. Challenges and Opportunities

All city leaders operate within a unique context. The challenges and opportunities that the city faces are derived from the socio-economic conditions, demographics, population size, accessibility, infrastructure, local business types etc. These form the baseline, or 'building blocks'[7] that the city leaders have to work from. These baselines are continually moving, and forecasting is used to assess future challenges. For example, a report on 'Demographic Challenges for European Regions' cites that "demographic change might lead to further increases in social polarization in Europe"[8] in the next 10 years.

According to the report "All Our Futures" (a forecasting report from the UK Office of the Deputy Prime Minister, based on a literature review, trends analysis, a Delphi survey and futures events) "in 2015 many of the pressures of government will manifest most dramatically at a local level. More flexibility and responsiveness at a local level would significantly enhance governments' capability to meet the challenges successfully" [7]. It is becoming important then, that city leaders actively engage in seeking out and responding to the opportunities and challenges faced by their particular locality.

Static snapshots of understanding are of limited value. This is because of the dynamic nature of the problem; the solution itself will change the operating context. For example the 'information economy', which has been driven by technological advances, is itself a driver of change for a city's challenges and opportunities. In turn, the way a city chooses to respond (perhaps through investing in start-ups or providing free public wifi) will affect the operating context yet further. Here, there is a reinforcing loop between the 'challenges and opportunities' and the solutions that might be implemented. The implication of this is that, if cities are to stay effective and responsive to citizen needs, they must continually evaluate the operating context through analysis of its challenges and opportunities. This can be achieved through local, national or international schemes such as the Covenant of Mayors [9].

IV. Policy Goals

The quantified challenges and opportunities are the basic facts about a city's context although 'opportunities' may be more subjective. "Once facts are accepted the ethics discussion begins as to how facts can be used or the appropriateness of facts in any specific case" [10]. Political leaders must translate these facts into goals or aspirations of where they would like to be; they must translate the "is" and "is-nots" into "ought" and "ought-nots" [11]. So, the 'challenges and opportunities' from part III, influence the development of policy goals. This process is heavily value-laden and requires the facts to be filtered through a political and cultural value system.

There is, however, an ongoing debate about the role of ICT in challenging the political and cultural paradigms of city leadership. The opportunities created by the 'smart city' "may lead us to a more fundamental choice between a privatized government (in which most issues are dealt with according to commercial relationships and principles, with services paid for by clients) and traditional, public government (in which many services considered to be of public interest are provided to citizens and businesses according to a variety of criteria not necessarily linked to commercial considerations)" [12]. The political context will fundamentally change how presented facts are interpreted into policy goals. City leaders must take care to ensure that the ability of ICT to outsource city services does not dictate the political direction, but that instead, investment in ICT is derived from a sound articulation of political, social and cultural values.

The policy goals are created to address the challenges and opportunities in a given context, and can directly affect citizen quality of life, environmental impact and civic engagement. In their research, Oxendine et al. found that "policy approaches regarding information technology interact with civic activity to predict both general internet use among citizens, as well as citizens' likelihood to use the Internet to seek out information about local government and community activities" [13]. There is therefore a tangible role for local and national governments to create policy goals that can genuinely respond to local challenges and capitalize on opportunities.

V. Public Value

The concept of public value management has become increasingly important for city leaders in recent years. O'Flynn explains that "public managers have multiple goals which in addition to the achievement of performance targets, are more broadly concerned with aspects such as steering networks of providers in the quest for public value creation, creating and maintaining trust, and responding to collective preferences of the citizenry" [14]. The creation of public value is described by Baptista as the "raison-d'etre of government," [12] and is distinct from simple cost cutting. "Public value argues that public services are distinctive because they are characterised by claims of rights by citizens to services that have been authorized and funded through some democratic process" [13]. It is delivered when solutions implemented adequately respond to a city's challenges and opportunities, and contribute the achievement of policy goals. It covers a host of outcomes from the tangible examples such as 'reduced crime on busses' to the more intangible delivery of 'feeling of safety' or 'level of happiness' [15]. Kelly and Muers [16] claim that public value can be achieved through the:

- Delivery of high quality services
- Achievement of outcomes that are seen as desirable by the public (will depend on opportunities and challenges as well as culture and values)
- Trust in public institutions

Kearns argues that "(public value) can be used both as an aid to judgment by governments when deciding what activities to undertake and as a yardstick against which to assess government performance" [17]. In this way, the delivery of public value determines the types of solutions that are implemented as well as delivering on policy goals.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

VI. SOLUTIONS

In order to address the challenges and opportunities of the city and deliver public value, city leaders invest in solutions. These solutions are actionable long or short term projects or programmes that can take many forms including regulation, facilitation or procurement. These are exemplified by smart grid implementation projects, e-ticketing or facilitation programmes such as the iShed in Bristol, the Bristol Council initiative "to produce creative technology collaborations" [18].

ICT or 'smart' solutions that the governments invest in can have direct and indirect implications for the creation of 'public value'. Cresswell et al. identify three core mechanisms in the creation of public value through ICT [19]:

1. Direct service impacts (which "occur when IT is embedded directly in a service delivery process, generating service changes that enhance value to the citizen")
2. Indirect service impacts (which "occur when back office or infrastructure investments produces changes in a government business process")
3. Mixed service and environmental value impact (where "the new IT is also linked to changes in the environment and relationships between the direct beneficiaries and other entities in the public area")

VII. MODEL

The core themes derived from the literature review include: Opportunities and challenges, policy goals, public value, and solutions. These themes are distinct but related, and the conceptual model in Figure 1 interprets the relationships between them. The model highlights that a city's challenges and opportunities influence city leaders' policy goals as well as the types of solution they might employ. These solutions are intended to resolve the challenges and opportunities of the city, but also deliver a wider contribution to public value, which in turn also supports the delivery of policy goals.

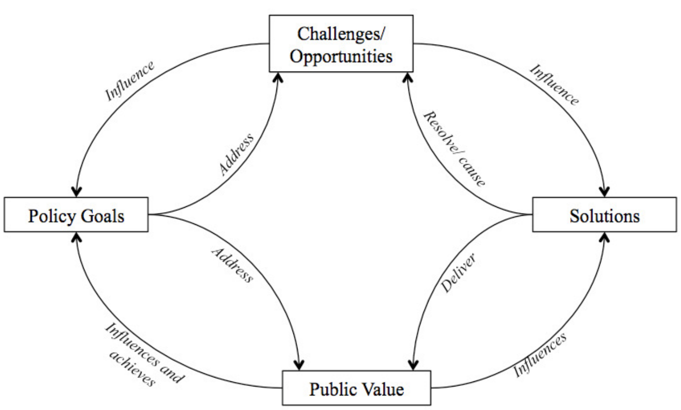


Figure 1. Grounded model of smart city policy and implementation concepts.

This model is not hierarchical. It is not the case that governments always set policy goals to directly address challenges and opportunities within their cities, for which appropriate solutions are derived. Often there is a 'solutions push,' i.e. a solution that is put in place because it seems to make sense and creates 'public value,' even if it does not directly come from a policy goal or a specifically quantified challenge. This model can be applied to interpreting the Greater London Authority's action around open data [20], which shows a city making an investment driven from a bottom-up understanding of creating public value. Emer Coleman saw that for a relatively small investment of £15,000 she could set up an Open Datastore for London. The decision to invest was driven by an inherent understanding of value to Londoners in terms of economic and social development, as well the delivery of better services and the transparency of government. It also responded to the city's 'challenges and opportunities' around private sector innovation and supporting start-up companies. Here, an investment in a solution was made as a contribution to public value rather than being instigated from a particular policy goal.

VIII. CONCLUSION AND FURTHER WORK

The model in figure 1 shows two core influencing features which are; the "challenges and opportunities" and the concept of "public value". These two factors form the context for any government action, and should be specifically articulated and continually explored. To respond to these two factors, governments act in two ways. Firstly they create policy goals which are intended to set investment priorities. These policy goals are explicitly or implicitly responding to local challenges and opportunities, as well as attempting to create public value. Secondly, city leaders intervene with solutions. Likewise, these respond to the core influencers in the system, and, through the tackling of these influencers, are intended to achieve policy goals.

Achieving policy goals through implementing solutions is often oversimplified, with public sector investment being directly informed by policy goals without consideration of the wider context. This is exemplified by the multiple failures of bike hire schemes across the UK (e.g., in Reading, Cardiff and Bristol). Often these schemes are developed directly from a local government policy to increase healthy lifestyles and reduce carbon emissions, but fail to understand where the true public value lies. This might include, for example, the failure to understand primary user groups and usage patterns. The ultimate goal should be to align policy goals with appropriate solutions. However, this is not achieved through directly jumping between them, but instead requires a consideration of the wider view offered by this model.

Achieving this in the complex city environment where long and short term local goals must be balanced with national and international-interest policies is especially demanding. For city leaders, this complexity is further compounded by the huge advancements in the ICT industry, which has fundamentally shifted the citizen behavior, expectations, and the economy. Devising solution programmes that will work effectively within this system, as well as respond to global calls for emissions reductions will be the defining feature in local governance in the next ten years.

This simple model in its current state is unable to offer the level of insight that a strategic decision maker requires in order to justify investment. To develop this model further we plan to explore each of the core components in turn and create a more detailed, dynamic model. This could include causal mapping and system dynamics modeling. For example, to investigate how cities might leverage policy goals and solutions to maximize public value from the information economy, we first need to gain a comprehension of the value that is already provided by information (or perhaps 'information products' [4]) within the city. Undertaking this research goes some way to quantifying the 'challenges and opportunities' that exist within a certain city. This research will be undertaken using case studies, for which a 'value chain' approach will be adopted to map the source of, and quantify the value. This builds upon work by Mulligan in her book "The Communications Industries in the Era of Convergence" [5]. This understanding will be combined with an investigation into what is considered of 'public value' to that city. Combining these two will support the development of appropriate policy goals and, a solution roadmap. The intention is to support cities in maximizing the value created from the 'information economy', or 'smart city', and ensuring that they have a robust understanding of both the direction of their policy goals, and the appropriate solutions.

ACKNOWLEDGMENT

This work was supported by the Systems Centre at the University of Bristol, UK; the EPSRC funded Industrial Doctorate Centre in Systems (Grant EP/G037353/1); and Arup. The authors would like to thank Volker Buscher and John Davis for supporting this research.

REFERENCES

[1] Kenna, M. "Social media: following EU public diplomacy and friending MENA", European Policy Centre, Policy brief, July 2011.

[2] Bright, P. "How the London riots showed us two sides of social networking", available at: < http://arstechnica.com/techpolicy/news/2011/08/the-two-sides-of-social-networking-ondisplay-in-the-london-riots.ars >, September 2011. [retrieved: March, 2012].

[3] Kazman, R. and Chen, H. The metropolis model: a new logic for development of corwdsourced systems, Communications of The ACM, 2009.

[4] Webb, M et al. "Information Marketplaces: the new economics of cities", The Climate Group, Arup, Accenture, Horizon, London, December 2011.

[5] Mulligan, C. "The Communications Industries in the Era of Convergence", Routledge, 2011.

[6] Sevtsuk, A. and Beinart, J. The effects of ICT on City Form. MIT School of Architecture and Planning, Cambridge, 2005.

[7] Tavistock Institute, "All Our Futures: The challenges for local governance in 2015", The Office of the deputy prime minister, London, 2006.

[8] Fésüs, G., Rillaers, A., Poelman, H., and Gákova Z. "Regions 2020: Demographic challenges for European regions" Brussels, November 2008.

[9] The Covenant of mayors, available at < http://www.eumayors.eu/index_en.html > , Jan 2012. [retrieved: March, 2012].

[10] Gortner, H. "Values and Ethics", 2001. In: Cooper, T. "Handbook of administrative ethics", Marcel Dekker Inc. New York, 2001.

[11] Hume, D. "A Treatise of Human Nature", OUP Oxford, 1978.

[12] Baptista, M. "e-Government and State Reform: Policy Dilemmas for Europe". the Electronic Journal of e-Government, 2005.

[13] Oxendine, A., Sullivan, J., Borgida, E., Riedel, E., Jackson, M., and Dial, J. The Importance of Political Context for Understanding Civic Engagement: A Longitudinal Analysis. Ploitical Behaviour, Vol 29, No1. (mar 2007) pp. 31-67.

[14] O'Flynn, J. 'From New Public Management to Public Value: Paradigmatic Change and Managerial Implications', The Australian Journal of Public Administration, 2007, vol 66 no 3, pp. 353-366.

[15] Rodríguez-Posea, A. and Maslauskaitec, K. "Can policy make us happier? Individual characteristics, socio-economic factors and life satisfaction in Central and Eastern Europe". Cambridge Journal of Regions, Economy and society, 2011.

[16] Kelly, G. and Muers, S. Creating Public Value, Cabinet Office Strategy Unit, 2002.

[17] Kearns, I. "Public Value and eGovernment", Institute for Public Policy Research, London, 2004.

[18] Ished, avialable at < http://www.watershed.co.uk/ished/ > 2012. [retrieved: March, 2012].

[19] Cresswell, A., Burke, G., and Pardo, T. "Advancing Return on Investment Analysis for Government IT: A Public Value Framework".

[20] London Datastore, available at < http://data.london.gov.uk/ > 2012. [retrieved: March, 2012].

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

Human Behavior Considerations in Metrics for Smart Infrastructures

Krishna Kant
George Mason University
kkant@gmu.edu

Abstract—**Smart systems such as physical infrastructures infused with information technology offer a great opportunity for satisfying the customer needs and managing the resources more effectively and efficiently. In this paper, we examine the smart infrastructures in terms of their sustainability potential, which in turn requires a consideration of human aspects. We consider the currently defined metrics for the smart grid and point out the lack of metrics to quantify this aspect. We then present a simple model for smart grid that considers the carbon footprint and energy consumption that would result by accounting for certain aspects of human behavior. Such a model is useful in assessing whether the sustainability objectives are being met and what actions are required in this regard.**

Keywords: Smart Infrastructure, Sustainability, Rebound effect, Carbon footprint, human behavior

I. Introduction

The world we live in is experiencing rapid and likely long-term changes that are already posing difficult sustainability challenges. For example, the increasing population and ongoing migration of populations to mega-cities puts a severe stress on the built infrastructure and the environment. The ongoing efforts to embed information technology into the built infrastructure to make it smart provides a unique opportunity to address this sustainability challenge by promoting efficient use of resources and providing critical services that enhance safety and human well-being. Thus, an important metric to consider for smart infrastructure is sustainability, which requires a thorough understanding of the interactions between built, natural and social systems [9]. In this paper, we would like to make a small beginning by considering the impact of human behavior on the resource usage and the corresponding carbon footprint. We illustrate the considerations involved in doing so, and then propose a simple metric for capturing the resource use efficiency in smart infrastructures.

The organization of the paper is as follows. Section II discusses the role of smart infrastructure in sustainability and points out the need for modeling human behavior. Section III takes the example of smart-grid where metrics are well developed and points out their deficiencies in capturing the human behavior. Section IV then identifies the basic issues that need to be considered in adequately considering the long-term sustainability aspects for smart infrastructures. Section V develops simple equations for characterizing the impact of human behavior on energy consumption and carbon footprint of smart grid and shows some sample results. Finally, section VI concludes the paper.

II. Sustainability and Smart Infrastructure

It is well recognized that the legacy infrastructure, including power grid, transportation infrastructure, vehicles, homes, buildings, factories, etc. can be quite wasteful of resources, and an embedding of intelligence in form of active monitoring, management and coordination of distributed resources could substantially increase the efficiencies. For example, Smart grid enables the integration of many disparate sources of energy from rooftop solar in a home to large electric utility installations. Smart grid can also keep track of energy consumption profiles and thereby direct energy where it is needed and in the process reduce peak generation and storage requirements and hence the ultimate cost to the consumers.

Although the efficient energy use and reduction of waste contribute to sustainability, the smart-grid by itself does not provide any mechanism to reduce consumption of energy. Smart grid only enhances flexibility and lowers cost for the consumer (in terms of avoidance of local storage and ability to sell excess energy back to the utility). Following the well known *Jevon's paradox*[6], or the *rebound effect*, the enhanced flexibility and lower cost will invariably spur the consumer to increase the energy consumption. In particular, if the lowered cost is ploughed back into additional energy purchase, the result could be net increase in peak energy draw, in plant size, and the carbon emissions. Thus, understanding human behavior is crucial to evaluating the long-term sustainability impact of the smart grid [3], [8], [13]. Understanding human behavior is even more important under disaster scenarios to avoid instabilities in the grid and make the most effective use of the surviving infrastructure – although we do not address this aspect in the paper.

Similar arguments hold for other smart infrastructures as well. A smart transportation system can provide congestion alerts, determine optimal routes, help avoid accidents and hazardous conditions, etc. The resultant reduction in fuel consumption and shorter trip to the destination can reduce carbon impact of the trip and make the roads much safer. Nevertheless, the additional infrastructure itself may contribute to additional carbon footprint; the increased travel demand prompted by reduced travel time, less congestion, and safer travel could have an even greater impact.

Smart buildings and homes provide an interesting example of infrastructure where the main objective of the intelligence is to reduce consumption of resources such as energy or water. While turning off cooling/heating/lighting in unoccupied areas

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

can surely reduce energy consumption substantially, this applies only to buildings with low occupancies and highly granular controls. In other instances, the automated control could well result in higher energy consumption than a conscious manual control by the occupants. Once again, understanding human behavior and accounting it in the evaluation is crucial for assessing the overall benefit.

III. Current Metrics and Their Deficiencies

Since smart grid has been advocated as a significant improvement over the existing and well entrenched power grid of yesterday, the issue of metrics to quantify the cost-benefit trade-off of the smart grid has been a topic of substantial current interest. In particular, the investment grants and demonstration projects funded by US Department of Energy (DOE) and Electric Power Research Institute (EPRI) have required evaluation of the projects along a variety of metrics that they have defined. In the following, we review these and related metrics and point out their deficiencies from the sustainability perspective. Similar observations can be made about other infrastructures as well.

The DOE analytical framework for smart-grid [1] establishes around 25 benefit metrics. These relate to costs, loading of different components, reliability, power quality, losses, power factor control, etc. A few of the metrics also relate to sustainability and concern capacity deferral and emission reduction due to smart energy management and integration of distributed renewable energy.

There are several other evaluation attempts for smart grid in the literature. In particular, Faruqui [4] shows the advantages of reduction in peak energy production, energy efficiency and distributed generation. Miller [10] provides a conceptual framework with multiple metrics under the broad categories of system efficiency, economic issues, reliability, security, environmental issues and safety. There are several other metric related studies, as summarized in the EPRI document [2]. Most of these metrics are effectively similar in nature, and they all lack consideration of long-term sustainability issues and a detailed understanding of the human behavior.

IV. Understanding Human Behavior

In understanding human behavior relevant to resource use, we need to address the following aspects of smart infrastructures: demand shaping mechanisms, extent of customer compliance with behavior monitoring, social influence, and behavior shifting over time. In the following subsections, we explain these and discuss how we plan to account for them. It is important to note here that although there is a rich literature to understand human behavior at a more detailed and elemental level (cultural, social, cognitive, and psychological) [11], our interests are at a higher level here. The more detailed models could presumably be used in deriving some of the parameters that we desired here.

A. Demand Shaping

Demand shaping refers to user's response to non-coercive mechanisms designed to influence the resource usage. (We only consider non-coercive mechanisms since the main issue with coercive mechanisms is compliance, which is covered in the next subsection.) The most common example of demand shaping is demand sensitive pricing. Another common example is to provide appropriate feedback to the user on its resource consumption. We discuss these in the following.

Demand shaping via dynamic pricing is a well known mechanism that is already being used to some extent in the current power grid. Studies show a 50% increase in price reduces household energy demand by only 11-15% and the tripling of price reduces consumption by 29-36% [13]. While pricing is an effective tool for smoothing out short-term demand variations, the rate regulations often come in the way of using it as a longer-term demand shaping mechanism. For example, while a utility can charge more during peak consumption periods, it often cannot raise the rate uniformly over long periods. As distributed generation takes hold, where the consumer generates its own energy, there is less role for pricing based behavior change.

Providing energy use feedback to user is a well known technique for influencing energy consumption behavior. It is important to note that feedback is not simply about providing information – the form in which the information is provided, the way it is presented, and how it is reinforced are crucial to making an impact on the energy consumption. The granularity of feedback is also critical – making a customer aware of using too much energy on hot water is far more useful and actionable than simply providing the overall increase in energy usage. It is also important to note that a feedback with the intent of decreasing energy consumption could sometimes have an opposite impact. This happens when the user determines that its consumption is lower than its peers or the effort to reduce consumption did not really make much difference to the energy bill. Experiments show that feedback can result in change in energy consumption from -5% to +20% [3].

Voluntary demand shaping depends on the choices made by the user, and an understanding of how users react to such mechanisms can improve predictability and ultimately the design of the mechanisms. Unfortunately, humans cannot be modeled as rational actors that simply maximize their utilities, even though a lot of pricing driven models assume this to be the case [12]. Users are often influenced by a host of factors, many of which may be ephemeral. For example, the adoption of an energy conservation idea or technique by an individual could depend on such diverse things as mood or emotional state, the way the idea is presented, prevalent social and cultural norms, ease of understanding the benefits, novelty, etc. [13].

B. User Compliance

The operation of a smart infrastructure involves collection of a variety of data about the way customers use it and often generates advisory data to enable them to make better choices. User compliance in terms of both accurately providing the requested data (input side) and following the advisory (output side) is crucial for smooth operation and resource efficiency of the entire system.

Let use start with the input side. A significant advantage over of smart grid over the current power grid is the active monitoring of the energy demand that allows a better management of the flow of energy and hence less conservative sizing of the facilities. However, the key question is whether the customer would feel comfortable with the active monitoring and what kind of data he/she will allow to be collected? The granularity of data collection acceptable to the user affects the advantage that can be derived from it. For example, if the customer is unwilling to allow individual monitoring of major energy consuming systems in a home/building, it is not possible to generate customized strategies for reducing the energy consumption. Similar comments apply to other smart infrastructures. In the smart transportation system context, if the drivers are unwilling to share their destinations and planned routes, the system cannot derive very accurate information about expected delays and future congestion.

There is a similar compliance issue on the output side. If a smart transportation system provides drivers feedback regarding the desirable routes or routes to avoid, but such feedback is ignored, the system will not be able to do a good job of managing delays and congestion irrespective of how much information it collects or how good an analysis job it does. In fact, erratic compliance could occasionally lead to traffic and congestion that is much worse than the one without any intelligent monitoring. Similarly, if the smart grid generates sound advice on how to smooth out the energy consumption profile w/o sacrificing comfort, and the advice is poorly used, it could lead to worse supply-demand mismatches.

Another area where compliance is an issue are legal mechanisms that prohibit wasteful use or mandate frugal use of resources. However, such mechanisms are normally viable only in cases of extreme resource shortages which we do not consider here.

C. Social Influence

An individual's behavior is often driven by the behavior of others around it, either in form of imitation/conformance or as a contrarian. For example, a driver that sees a lot of other people exiting the road may either also decide to exit (conformance or herd behavior) or decide to stay (to potentially benefit from the herd behavior exhibited by others). Similarly, people may imitate or contradict other people in their neighborhood in terms of energy or water use.

The characterization of social influence is itself a challenging problem but has been studied extensively particularly in the context of "herding behavior" that is routinely seen in complex adaptive systems such as investing, driving/walking, emergency evacuation, etc. [14]. As with other parameters, devising general model for it is extremely difficult. Even a direct measurement could be somewhat challenging since it may not be clear if the observed behavior represents herding or something else.

The social influence can be exploited to nudge people towards sustainability, though it could also prove ineffective. For example, if home-owner is simply provided data about the energy consumption of other similar homes in the neighborhood, the effect is likely to be conformance rather than reduction in energy consumption. In particular, those with high energy consumption may reduce it, but those on the lower end may increase their consumption. However, creative mechanisms for feedback and positive reinforcement could help lower the consumption on the higher end without raising the lower end.

D. Behavior Shifting

This aspect directly relates to the long-term shift in the resource usage that may be triggered by the smart infrastructure itself. The qualification here is essential since we do not want to include influence of other extraneous factors such as rising incomes of the population served by the smart infrastructure. Thus the primary mechanism for behavior change considered here is the rebound effect. This is likely to be mostly cost driven – the lower cost brought about by improved efficiencies and ability of customer to sell power to utilities could make them gradually increase their consumption. However, other factors may also play a role. For example, the "feel good" emotion about locally harvested renewable power may make people more comfortable with increasing their energy consumption. Generally, such behavior shifting will have a long time constant. It is generally very difficult to characterize since the impact depends on a large variety of poorly understood human behavior issues; therefore, we merely illustrate it here without making any claims regarding the definitiveness of our assumptions.

Wang [13] contains a wealth of information regarding the behavioral impact on energy consumption. It also quotes data from 20 different studies to quantify the magnitude of the rebound effect in the context of energy use for various purposes. The high variability in the estimates in these studies (not included here for brevity) reflects the difficulty in ascertaining the impact of human behavior.

E. Human Behavior Centric Efficiency Metric

In this section, we propose a new metric that accounts for the influence of the above factors with respect to resource consumption and hence sustainability. The goal of the metric is to quantify to what extent human factors can increase or decrease the resource consumption (or its carbon footprint). The metric thus provides an estimate of what can be achieved by putting more effort into influencing human behavior.

Let R_0 denote the baseline resource consumption and R_x the modified resource consumption due to influence x where x takes the following values:

- $x = pf$: Influence of private feedback to customer on his/her resource consumption. Here, "private" means that a user receives data on his own consumption, and not that of others.
- $x = si$: Influence of feedback that compares a customer's resource consumption against that of his/her peers.
- $x = bs$: Influence of long-term behavior shifting (primarily rebound effect).

The three factors above have been defined in such a way that they are reasonably independent of one another. It must be acknowledged, however, that seemingly independent human behavior factors could be related in complex ways; therefore, one cannot claim complete independence. For example, it is possible that the extent of behavior shift depends on how much resource reduction we start out with. Nevertheless, we assume independence for simplicity.

Let us now define the metric γ_x as R_x/R_0 which gives the fractional resource consumption relative to the baseline. Furthermore, by exploiting the independence assumption, we also define total metric γ_t as follows:

$$\gamma_t = \gamma_{pf}\gamma_{si}\gamma_{bs} \quad (1)$$

The combined metric γ_t allows us to assess the net advantage resulting from the behavioral aspects, with $\gamma < 1$ being desirable. All 3 individual factors, and hence γ_t will typically vary with time, and may or may not settle to a long-term value. For example, γ_{pf} may go down slowly over months as consumers begin to react to feedback and then settle, go down initially and then creep up (initial enthusiasm for conservation that wanes over time), or show other more complex behavior. Thus a periodic assessment of the γ_t metric may be in order, with the intent that whenever it tends to inch up significantly, new mechanisms can be tried to drive it down. From a modeling perspective, we may be interested in predicted range for γ_t over the long term either with or without corrective actions.

The influence of feedback can be attributed to 3 different factors: (a) The granularity of usage data provided by the customer, (b) the granularity of feedback provided to the customer, and (c) the customer compliance, which in turn depends on the quality and effectiveness of the feedback. We optimistically assume that (b) is not an issue; i.e., the system always attempts to generate as granular a feedback as the monitored data would allow. The user provided data granularity may be limited due to a variety of factors including technological, regulatory, or customer choice. Even the technological factors could be a significant barrier since putting meters or sensors at multiple points could be expensive or impractical. In this regard, a significant amount of work exists on "signature analysis" to identify components of resource consumption [5]; however, such measurements are prone to errors. In view of this, it is not useful to determine γ_{pf} as a function of some measure of granularity; instead, it suffices to estimate typical values of γ_{pf} under a couple of scenarios and best presentational practices.

The metrics γ_x and γ_t can also be defined relative to carbon footprint instead of resource consumption. We denote these alternate versions as Γ_x and Γ_t respectively. The main advantage of considering these is their ability to distinguish between various technologies in terms of sustainability (e.g., renewable vs. fossil fuel based energy conversion). Note that the carbon footprint comes not only from the direct use of the resource (e.g., use of gasoline in the car) but also from the machinery (e.g., the car itself) and the associated infrastructure/operations (e.g., extraction and transportation of materials and fuel). A comprehensive accounting of all components of the carbon footprint is beyond the scope of this paper. Instead, we assume that for a given technology, the carbon footprint can be related to the resource use via a known constant factor.

V. EFFICIENCY METRIC FOR SMART GRID

In this section we focus specifically on the smart grid and address the question of under what conditions does smart grid lead to decrease in the energy consumption and the carbon footprint? In particular, we compare the sustainability advantages of integrating distributed renewable power generation by customers into the smart-grid vs. the traditional model of central utility based power distributed to the customers.[1] Our analysis here is rather simple and is only meant to illustrate the usefulness of the efficiency metric defined above. It is possible to take this analysis much further, but that is beyond the scope of this paper.

Let us assume that a smart grid serves a community of N homogeneous customers, each of which consumes R_c units and generates renewable energy G_c, at the cost of η_c dollars/unit (or KWHr). Of this, a fraction α is sold to the utility company at the cost of η_{cu} dollars/unit and the rest is consumed locally. Note that α is intended to be simply a long-term average fraction. Depending on the variability of the demand, the actual energy exported to the utility will vary with time.

Let G_u denote the required generation capacity of the utility. We assume that G_u is just enough to satisfy the demands of all N customers if they do not generate any energy on their own. The generated amount also includes excess to account for the transmission and distribution (T&D) losses. Thus, if f_L is the T&D loss fraction, we have:

$$G_u = NR_c/(1 - f_L) \quad (2)$$

Let R'_c denote the energy consumption of each customer under distributed generation. If there is no behavior change, $R'_c = R_c$, otherwise, the two could be different. Let G'_u denote the central generation requirements in this case. It is easy to see that:

$$G'_u = N[R'_c - G_c(1 - \alpha f_L)]/(1 - f_L) \quad (3)$$

Here, $R'_c - G_c$ is the amount of energy that the customer needs from the utility. Furthermore, since the customer feeds αG_c energy to the utility (to be recycled back to same or different customer), a fraction f_L of this is lost and needs to be generated centrally. Finally, all power generated by the central utility must be further bumped up by the factor $1/(1 - f_L0)$ to account for forward T&D losses.

We assume that the utility company generates energy from non-renewable sources and sells it at the cost of η_{uc}. It is assumed that $\eta_{cu} < \eta_{uc}$, i.e., the customer receives less from utility than utility's energy selling rate. (Although currently there are instances where the utilities pay more to customers than their selling rate, this is an unsustainable model in the long run.) Generally, one would expect that $\eta_c > \eta_{uc}$, i.e., small distributed generation is more expensive than the centralized generation if the up-front, leasing or maintenance

[1] In what follows, distributed generation is used as a codeword to refer to smart-grid, whereas centralized-only generation is used to refer to the conventional grid.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

costs of the infrastructure are also taken into account. With centralized generation, the energy price paid by a customer, P_c is given by:

$$P_c = R_c \eta_{uc} \tag{4}$$

With distributed generation, it changes to:

$$P'_c = [R'_c - G_c(1 - \alpha f_L)]\eta_{uc} + G_c[\eta_c - \alpha(1 - f_L)\eta_{cu}] \tag{5}$$

where the first term represents the cost of buying energy from the utility, second term the cost of generating it locally, and third term is the discount for selling energy to the utility. Note that in the first term, the customer must buy back the energy it sells to the utility at the regular rate. Also, just like the customer, the utility pays for only the energy it sees on its end. Thus, even though the customer sells $G_c\alpha$, the utility only sees a fraction $(1 - f_L)$ of this.

It is important to note here that it is not so much the actual cost η_c, but rather the customer perceived cost, say η'_c that matters. As in many other contexts of do-it-yourself scenarios, the customer often tends to ignore the fixed initial investment and regular maintenance costs and only considers the running cost. For example, when considering the cost of driving vs. taking public transport, people often only consider the fuel cost of driving. With renewable energy such as solar where the running costs are essentially zero, η'_c can be substantially lower than η_c, perhaps even zero. In this case, the customer may perceive as gaining even without selling anything back to the utility. This aspect is crucial for the rebound effect discussed below.

Let us now estimate the γ_t factors for our simplified situation. We shall assume that the feedback can modestly lower the energy consumption as reported in [13]. Since we are considering an idealized situation where all N customers have equal energy consumption, social influence has little role to play here, and $\gamma_{si} = 1$.

In order to estimate γ_{bs}, we assume a limited rebound effect where the customer increases its energy consumption to match the cost associated with the case of centralized generation only (that is, any energy reduction effects due to γ_{pf} and γ_{si} factors still hold). Thus, by equating the costs from eqns 4 and 5 and setting $\gamma_{bs} = \frac{R'_c}{R_c}$, we have:

$$\gamma_{bs} == 1 + [\alpha(1 - f_L)\frac{\eta_{cu}}{\eta_{uc}} - \frac{\eta_c}{\eta_{uc}} + (1 - \alpha f_L)]\frac{G_c}{R_c} \tag{6}$$

In reality, there is a limit to how much energy a customer can consume. We assume that a customer cannot increase his consumption by more than by a factor denoted as "Limit". And finally, $\gamma_t = \gamma_{pf}\gamma_{bs}$.

Let us now consider the estimation of Γ_t factors. We denote the carbon footprint of resource R as $\zeta(R)$. As discussed before, we assume that $\zeta(R) = C\,R$ where C is an appropriate constant. In particular, we use two constants, C_r for renewable energy, and C_f for fossil fuel based energy. Both of these constants would generally depend on the size of the installation, with larger installations having a lower carbon footprint per watt. In all cases, the carbon footprint needs to consider the average energy generation per customer, not just the consumption. The difference – T&D losses – do contribute to carbon footprint. With this, and eqn (2), we have:

$$\zeta(R_c) = C_f\,G_u/N = C_f R_c/(1 - f_L) \tag{7}$$

In case of smart grid where a customer produces G_c KW of energy locally and requires generation of G'_u/N units centrally. Therefore, from eqn (3), we have

$$\begin{aligned} \zeta(R'_c) &= C_r\,G_c + C_f\,G'_u/N \\ &= C_r\,G_c + C_f\,[\gamma_{bs}R_c - G_c(1-\alpha f_L)]/(1-f_L) \end{aligned} \tag{8}$$

Since $\Gamma_{bs} = \zeta(R'_c)/\zeta(R_c)$, it follows that

$$\begin{aligned} \Gamma_{bs} &= \frac{[C_r\,G_c(1 - f_L) + C_f\,[\gamma_{bs}R_c - G_c(1 - \alpha f_L)]]}{C_f R_c} \\ &= \gamma_{bs} - G_c/R_c[1 - \alpha f_L - C_r/C_f(1 - f_L)] \end{aligned} \tag{9}$$

Also, $\Gamma_t = \Gamma_{pf}\Gamma_{bs}$.

We now show the behavior of γ_t and Γ_t as a function of various parameters. For this, let us first describe the situation considered in our rather simple model. We assume that the customer receives power from the utility (η_{uc}) at the rate of 12 cents/KWHr, and the utility is willing to purchase power from the customer (η_{cu}) at 9 cents/KWHr. We assume that the real cost of locally generated power is 15 cents/KWHr, but as indicated earlier, the customer may perceive a lower cost. Since our simple analysis does not use the underlying costs, the results are the same irrespective of whether the assumed costs are real or perceived. Therefore, we will show results for (real or perceived) rates of 0, 5, 10, and 15 cents/KWHr. We assume that the T&D losses are 20%, which are somewhat on the optimistic side but perhaps achievable with smart grid. Finally, we assume that the smart grid can maintain a 5% gain in efficiency of the regular grid due to continuous customer feedback. We also assume that the rebound effect is limited to at most 60% increase in energy consumption over the baseline; i.e., the "Limit" parameter is set to 1.6.

We assume that a customer generates some percentage of its requirements G_c/R_c, locally and receives rest from the grid. If the local generation is small (up to $\theta_0 = 20\%$ in our example), all power can be used locally and none is sold to the utility. At higher local generation capacity, we assume that the customer is able to absorb $\theta_1 = 25\%$ of the excess, and sells the rest. That is,

$$\alpha = \begin{cases} 0 & \text{if } G_c/R_c \le \theta_0 \\ \theta_1(G_c/R_c - \theta_0) & \text{otherwise} \end{cases} \tag{10}$$

The parameter θ_0 depends on customer's decision regarding the sizing of its local generation capability and θ_1 indicates variability of his demand. If the demand is constant, we can set $\theta_1 = 1$; that is, the customer will use all of the local power up to the limit θ_0 and then export all the rest. A local power storage would allow the customer to reduce θ_1 so that it can keep the local power to handle its short term demand variations.

For the carbon footprint, we assume that the C_f is relatively independent of the capacity, but C_r decreases modestly with the size of the installation. In particular, we assume that as a function of G_c/R_c, C_r/C_f varies from 0.25 for a small installation to 0.15 for the case of $G_c/R_c = 1$. That is, we are assuming that a large renewable plant has only about 15% carbon footprint as compared to a conventional plant.

Figs. 1 and 2 show the energy and carbon footprint efficiencies of smart grid (γ_t and Γ_t) as a function of G_c/R_c

and for different costs η_c. Let us first compare the different curves. As expected, if the perceived or real cost is lower, the rebound effect prompts the user to consume more energy (up to 60%), which explains the systematic ordering of the 4 curves. In particular, if the customer thinks that its local energy costs significantly less than the utility power, the energy consumption goes higher than that for the centralized generation case (i.e., more than 1.0). The carbon footprint shows a similar but somewhat better behavior, as we shall discuss shortly.

Let us now consider the behavior with respect to G_c/R_c. When the perceived or real cost of local generation is small, an increasing G_c implies that the user is able to reduce its cost which in turn prompts him to consume more energy. Therefore, the consumption also increases monotonically with G_c (up to the assumed limit). Note that if the perceive or real cost of consumer generated power is higher than the cost of utility power, the net effect is the price based reduction in consumption as shown by the $\eta_c = 0.15$ curve. Not surprisingly, when the local generation fraction $G_c = 0$, all curves merge at the same point. This point is really at 0.95, rather than 1.0, and reflects the 5% feedback based advantage of the smart grid.

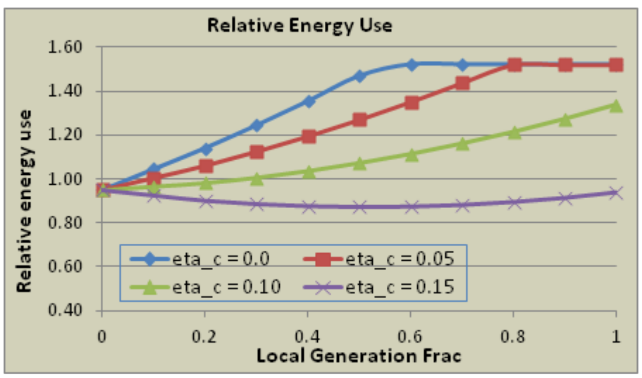


Fig. 1. Relative Energy Use vs. locally generated power fraction

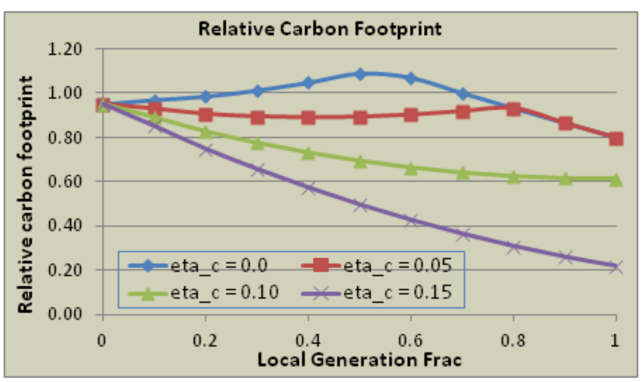


Fig. 2. Relative Carbon Footprint vs. locally generated power fraction

The primary difference between carbon footprint and energy use curves is our assumption that the local generation by the customer uses renewable source which has a much lower carbon footprint than the utility generated power. This aspect works against the increase in carbon footprint due to increase in energy consumption, hence the rather peculiar shape of the curves. In particular, when η_c is closer to the actual cost, the carbon footprint actually shows a monotonic decline with G_c/R_c.

It can be seen that in spite of its simplicity, the model can be used to address a number of scenarios. The model can also be tweaked to study a time-series of dynamically varying demands, though we do not pursue this due to lack of space.

VI. Conclusions

In this paper, we highlighted the need for metrics that capture the human behavior in characterizing the resource use and environmental impact of the smart infrastructure. We also developed a simple metric for the smart grid and showed its usefulness in evaluating the sustainability impact of the smart-grid. Similar analysis can be carried out for other smart infrastructures as well.

Admittedly, our model and analysis are rather simple, and the calibration parameters are somewhat arbitrarily chosen. The future work involves a more comprehensive modeling that accounts for heterogeneous customer base, a deeper modeling of the human behavior aspects, and the consideration of temporal aspect (i.e., the fact that energy demand and even the customer behavior varies with time). The ultimate worth of this exercise is to find ways of influencing human behavior in ways that can lead to more efficient resource usage and reducing its carbon impact without sacrificing comfort or safety. It is hoped that a more detailed modeling will provide insights into how to do this.

References

[1] S. J. Bossart and J. E. Bean, “Metrics and Benefits Analysis and Challenges for Smart Grid Field Projects”, U.S. Department of Energy, Proc. of EnergyTech, May 2011.
[2] Electric Power Research Institute, “Methodological Approach for Estimating the Benefits and Costs of Smart Grid Demonstration Projects”, Palo Alto, CA: 2010. 1020342.
[3] C. Fischer, “Feedback on household electricity consumption: a tool for saving energy”, Energy Efficiency, 2008, pp 79-104.
[4] A. Faruqui, R. Hledik, C. Davis, “Sizing up the Smart Grid,” presented at Elster EnergyAxis User Conference, Feb. 24, 2009
[5] J. Froehlich, E. Larson, S. Gupta, et al., “Disaggregated End-Use Energy Sensing for the Smart Grid”, IEEE Pervasive Computing, Vol.10, No.1, pp 28-39, Jan.-March 2011
[6] J.M. Polimeni, K. Mayumi, and M. Giampietro, *The Jevons Paradox and the Myth of Resource Efficiency Improvements*, Earthscan. pp 79140.
[7] T. Godfrey and C. Rodine, “Unified Metrics for Management of Smart Grid Home Area Networks”,
[8] E.D. Kuligowski and S.M. V. Gwynne, “The Need for Behavioral Theory in Evacuation Modeling”, Pedestrian and Evacuation Dynamics 2008, 2010, Part 3, 721-732.
[9] S.A. Levine and W.C. Clark, “Towards a science of sustainability”, Report from workshop, Nov 2009, available at http://www.nsf.gov/mps/dms/documents/SustainabilityWorkshopReport.pdf
[10] J. Miller, “Smart Grid Metrics: Monitoring Our Progress,” presented at the Smart Grid Implementation Workshop, June 19, 2008
[11] J. Thgersen and A. Grnhj, “Electricity saving in householdsA social cognitive approach”, Energy Policy, Vol. 38, No. 12, Dec 2010, pp7732-7743
[12] C. Triki and A. Violi, Dynamic pricing of electricty in retail markets, Q. J. Oper. Res., vol. 7, no. 1, pp 2136, Mar. 2009.
[13] J.H. Wang, “Behavioral Policy Modeling: Consumer Behavior Impacts on Residential Energy Consumption”, available at www.spp.gatech.edu/faculty/WOPRpapers/Wang.WOPR11.pdf
[14] L. Zhaoa, G. Yanga, W. Wanga, et al., “Herd behavior in a complex adaptive system”, Proc. of National Academy of Sciences, Vol 108, No 37, pp 15058-15063, Sept 2011.

Ambient Assistance Using Mobile Agents

Amine Arezki
Université de Versailles Saint-Quentin-en-Yvelines
Laboratoire d'Ingénierie des Systèmes de Versailles
Versailles, France
amine.arezki@lisv.uvsq.fr

Eric Monacelli and Yasser Alayli
Université de Versailles Saint-Quentin-en-Yvelines
Laboratoire d'Ingénierie des Systèmes de Versailles
Versailles, France
eric.monacelli@lisv.uvsq.fr, yasser.alayli@lisv.uvsq.fr

***Abstract*—In this paper, a method is presented to assure ambient assistance in an urban environment, using a mobile agent. The goal is to anticipate the assistance if needed. Therefore, the robot has to understand the human behaviour and a person's needs. We will determine how to focus on a moving subject and using interaction for confirmation, in order to provide an assistance if needed. Therefore, the trajectory type concept is used to define the first step of analyses, which is called the approach step. Combining this step with the field information provided by the mobile agent will insure a certain type of assistance. In terms of observation, two different views are employed to detect assistance requirements, *i.e.,* the Fix Intelligent Device and the Ambient Intelligent Devices, both communicating by Wi-Fi. The Fix Intelligent Device is a fix camera standing on a very top view allowing the detection of possible motions and classification of trajectories, using neural network. In our research, a Touch Ambient Intelligent Device is utilized, a mobile robot with three degrees of freedom, including a 3D camera (*Kinect*^TM^) and a touch screen to interact with the subject. In this paper, the behavior of the mobile agent, receiving detected trajectories emitted by the Fix Intelligent Device, is presented and it is shown that the human intervention is only needed in critical cases.**

Keywords-Human Activity Recognition; Ambient Assistance; Human Tracking; Robot Agent.

I. INTRODUCTION

The actual increased longevity in Europe and other developed countries [1] results in a fast growing population of senior citizens. Based on estimations [2], one-fifth of the elderly citizens will be 80 years and older by the middle of the century. Many elderly need nursing support to perform their normal daily activities, mostly due to the lack of strength or balance. Assistance robots can play an important role to reduce the workload of nurseries and increase the independence of elderly to perform their daily activities with little help from the people around them. At the same time, assistance robots developed for frail and disabled adults can be used for people with minor disabilities, as well.

That is one of the reasons that assistance robots have become one of the most rapidly developing fields in robotics. Among the assistance devices, there are some providing cognitive assistance and some increasing the mobility of the user.

A simple type of ambient assistance fully relying on human already exists. However, the disadvantages of this system are high costs, discontinuity and interpretation mistakes, such as animal detection, shadow, light, *etc*. A more complex type is using an automatic system (Vision, Sensors, *etc*.), employing the same disadvantages of interpretation mistakes eventually causing dangerous situations. In this work, the person's needs are interpreted, in order to provide assistance. The assistance robots described by Annicchiarico *et al.* [3] and Nagai *et al.* [4] are examples that can be used in displacement and transfer between two seated positions, respectively. Both examples are of relatively simple structure providing only a single type of assistance. While Chugo *et al.* reported a robot providing assistance for standing, walking and seating [5], Méderic *et al.* described ambient assistance for walking and the sit-to-stand transfer [6]. Assistance devices offering multi-support are normally large and of relatively complicated structure. Multi-robot assistance architecture has been proposed for frail and disabled adults in our laboratory [7] [8].

Our proposed system is designed for public places like waiting halls of train stations or airports or private places like apartments or offices.

The most challenging task is to understand the human behaviour, in order to understand a person's needs. In our research work, this detection is done by fusion of two different views to provide better results by having a continuous and safe automatic system. A robotic agent is used to confirm unclear information, *e.g.,* the doubt if the subject is a human, or only a shadow artefact movement. Therefore, our system proposes a solution to use humans only in critical cases.

In this paper, details of our system architecture are given first by explaining the role of the Touch Ambient Intelligent Device (TAID), followed by trajectory detections done by the Fix Intelligent Device (FID). The results of the combination of these two views (TAID and FID) are described subsequently. The experimental part with the results completes this paper, followed by the conclusion of our research.

II. AMBIENT ASSISTANCE USING MOBILE AGENT

The architecture shown in Figure 1 allows to visualize simultaneously several movements of subjects, as well as to analyse their trajectories.

The trajectory detection is done by image capture at first. The assisting robot TAID has been developed for this project to act on the field (hospital, station, airport, *etc.*), when needed, or to confirm information.

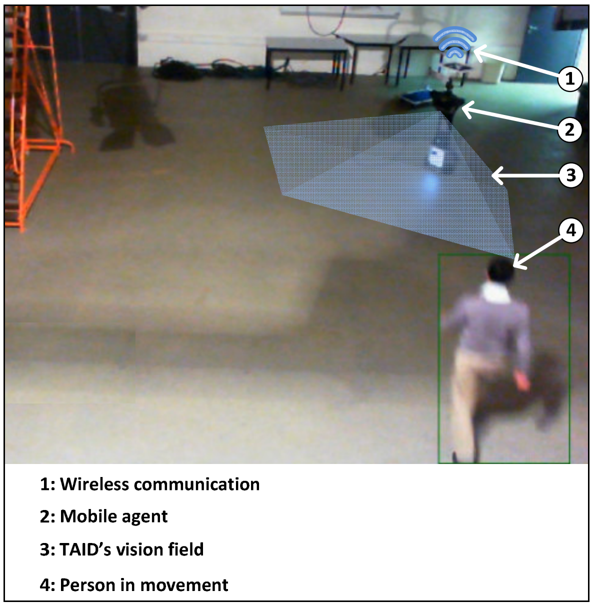


Figure 1. General view of the concept.

TAID is a robot with three degrees of freedom, allowing movement on X and Y axis, as well as rotation on Z axis (Figure 2). TAID is capable of lateral crablike movements, playing with direction and velocity of the wheel rotation.

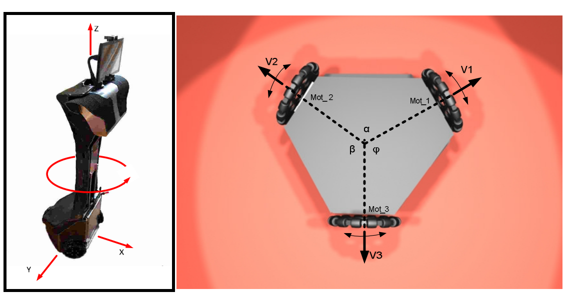


Figure 2. Degrees of freedom of TAID (on the left); Rotation axes of TAID's actions (on the right).

A robotic 3D sensor *Kinect*^{TM} of *Microsoft*^{TM} [9] has been embedded for simplification, allowing the detection and consequently the avoidance of obstacles, on top of the detection of subjects to determine their pose and to differentiate between human and animals.

A. *Robotic Assistance Description*

The communication between TAID and the subject is done by visual (*Vi*), tactile (*Ta*), vocal (*Vo*) or text interaction (*Te*).

Examples of interaction:

1. Subject: Making a hand signal → *Vi*.
2. TAID: Approaching and asking "DO YOU NEED HELP? RAISE YOUR RIGHT HAND FOR *YES* AND YOUR LEFT HAND FOR *NO*!" → *Te* and *Vo*.
3. Subject: Answering by giving the corresponding hand signal → *Vi* or by pressing the button *YES* → *Ta*.
4. TAID: Asking "WHAT INFORMATION DO YOU NEED?" → *Te* and *Vo*.

On the part of the robot, the question is always asked vocally, as well as by text using the robots graphic interface. The subject's response is then analysed by reading its gesture: Right hand raised for *YES* and left hand raised for *NO*. In order to avoid any incomprehension, vocal recognition was evaded, as the system is generally used in a noisy environment.

A quadrilateral transformation needs to be done primarily before a follow-up of any movement may start. The location of the camera has to permit the visualization of the platform to analyse.

B. *Movement Detection*

There are several approaches in a flux video for movement detection, based either on comparison of the current image video with one of the previous images (1), most commonly used, or with the background (2).

If a colour image is considered as example, a copy of the current image is done in grey scale, as well as of the previous image video. First of all, the region is determined by subtraction of the two different images. An image of white pixels is obtained at the place where the image $I(t_x)$ differs from the image $I(t_{x-1})$. If the surface is bigger than the defined lower limit, a movement alert is obtained. The inconvenient of this approach lies in the velocity of the object: If the object moves at a slow rate, the comparison of the current image with the previous image may provide insufficient results for object detection.

The comparison of the image $I(t_0)$ with the initial image $I(t_{in})$ allows the contour detection of a subject in movement, independently of the movement velocity. The inconvenient of this approach lies in the disappearance of the object: If an object is present during the first image capture of the background, but absent in the second image capture, the object is still detected at this place.

In our research work, a combination of these two approaches has been chosen, *i.e.,* the comparison with the

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

image $I(t_{x-100})$. The advantage of this approach lies in the movement detection, even when the subject is moving at a slow rate and even when the reference image is still being up-dated. Thus, the problem of disappearing objects from the background is solved, as well as their permanent detection.

C. *Robot Interaction*

In the following section, the different interactions of the robot are presented.

1. Hand Signal

The agent may capture a hand signal and hence, may directly orient itself towards the subject for interaction. The interaction is done by vocal and written questioning/answering.

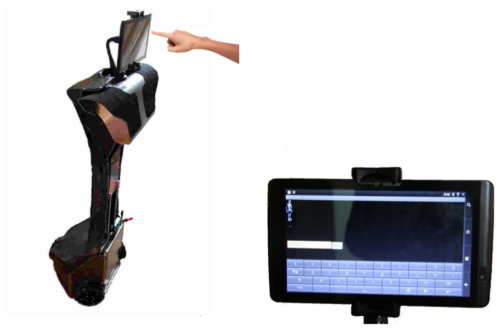


Figure 3. TAID's tactile interaction tool.

The mobile agent is using a tactile interaction tool, as shown in Figure 3, simplifying the interaction even in crowded and noisy environments.

2. Body Pose

The assisting agent TAID may detect a sitting or an on the ground lying subject, employing the 3D camera *Kinect*^{TM}.

The following two situations may occur, causing different reactions on the part of TAID :

- The agent detects the subject's falling down (subject standing and subsequently lying). In this case, the robot TAID will send an emergency alert.
- The agent detects a lying subject without knowing if the subject fell down or lied down on purpose. In this case, the robot TAID will interact with the subject first, asking him to confirm the need of assistance. For safety reasons, the robot will send an emergency alert after a certain time if the subject is not responding.

III. TRAJECTORY PATTERN

Five types of trajectories have been selected: Zigzag trajectory (A), circular trajectory (B), direct trajectory (C), back-and-forth trajectory (D) and random trajectory.

A. *Zigzag Trajectory*

A zigzag trajectory (T_Z) defines a hesitating or lost subject that does not know where to go. Hence, the goal is to assist the subject to regain orientation.

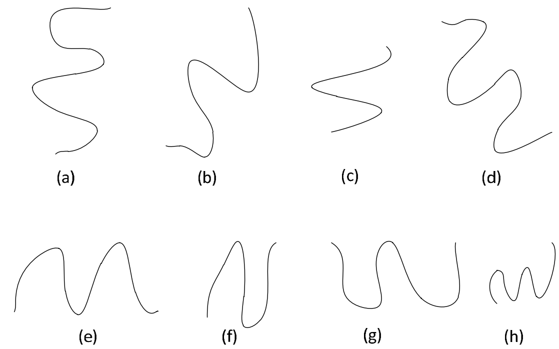


Figure 4. Examples of zigzag trajectories (T_Z).

Different T_z are possible, as it can be seen in Figure 4.

B. *Circular Trajectory*

A circular trajectory (T_c) can be a trajectory of a lost or waiting subject in movement, requiring assistance.

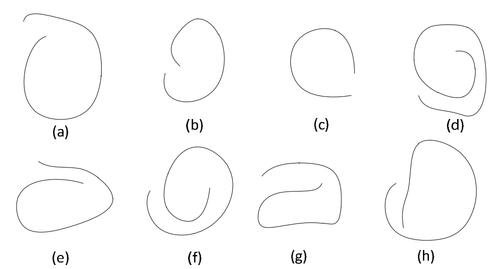


Figure 5. Examples of circular trajectories (T_c).

Figure 5 is representing some examples of such possible T_c.

C. *Direct Trajectory*

The direct trajectory (T_d) describes a decided subject, knowing where to go. The different possible paths are shown in Figure 6.

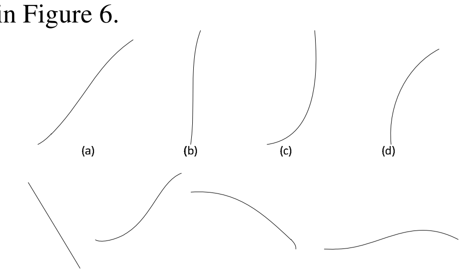


Figure 6. Examples of direct trajectories (T_d).

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

D. Back-and-Forth Trajectory

A back-and-forth trajectory (T_b) helps to anticipate the agent's position, in order to have a fast intervention if a zigzag or a circle trajectory is detected (Figure 7).

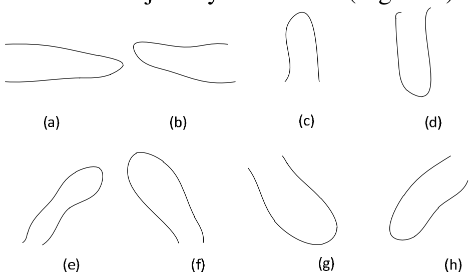


Figure 7. Examples of back-and-forth trajectories (T_b).

1) Trajectory detection

The detection of the trajectory type has been done by the use of a neural network [10] with 1000 inputs for each type of trajectory (Figure 8), allowing their classification.

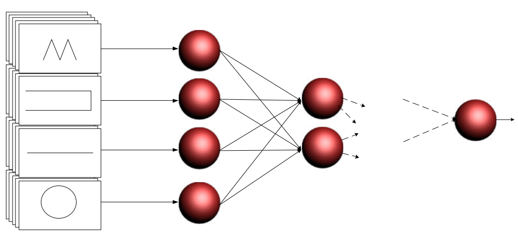


Figure 8. Illustration of the neural network model.

The fixed camera is connected to a computer *via* USB port, containing the algorithm of analysis and decision. The algorithm will transmit the command to the mobile robot TAID by Wi-Fi, using the User Datagram Protocol.

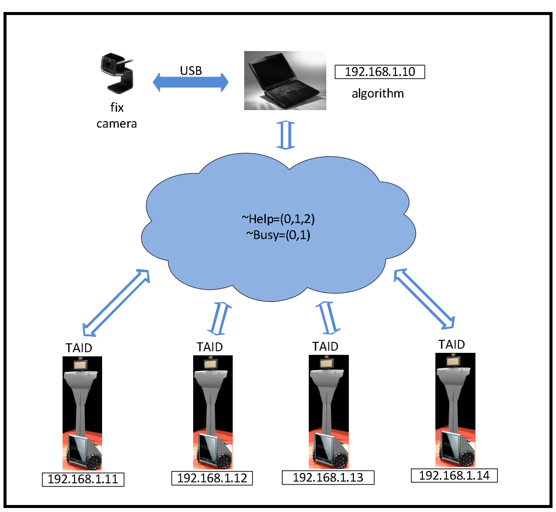


Figure 9. Illustration of the agent's communications.

Each mobile agent possesses its own IP address, as it is shown in Figure 9.

IV. SEMANTIC MODEL

During the agent movement, the FID is orienting the agent towards the circulating subject to have the possibility of information capture, emitted from the subject. The agents are able to detect following situations:

- A hand signal is coming from the subject [11]; this detection provokes TAID to approach the subject and to interact with the same.
- The subject is sitting or lying on the ground [11]; this detection provokes an emergency alert provided by TAID.

Figure 10 shows TAID's movement and its anticipation during analysis of the captured trajectory.

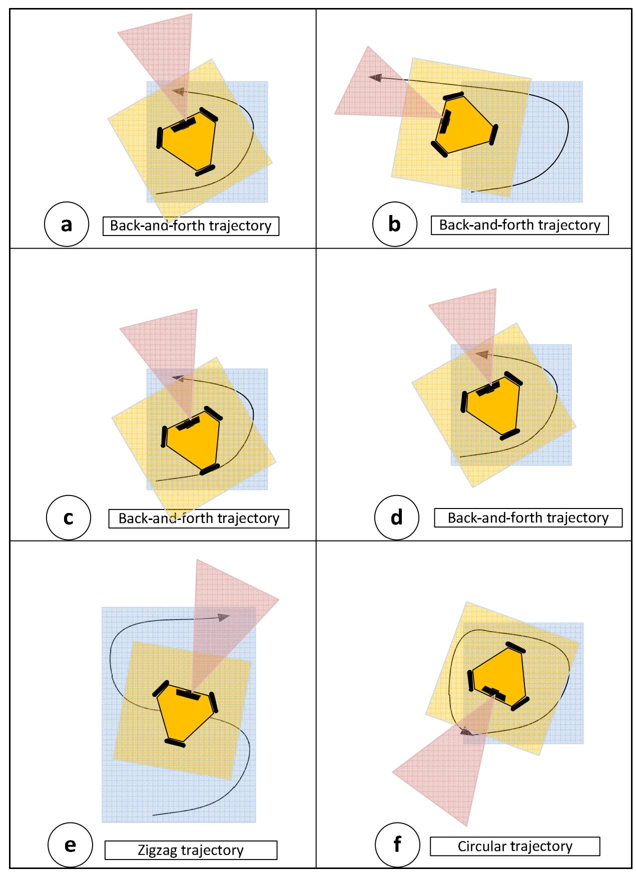


Figure 10. Examples of TAID's movement, based on the captured trajectory: (a-d) Back-and-forth trajectories; (e) Zigzag trajectory; (f) Circular trajectory.

V. EXPERIMENTAL SECTION

A volunteer was required to walk around in the test area, respecting certain trajectories and repeating these trajectories 9 times. Figure 11 illustrates the resulted trajectories.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

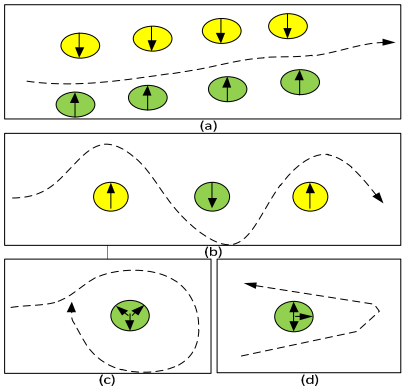


Figure 11. Illustration of the tested trajectories: (a) Direct trajectory ; (b) Zigzag trajectory; (c) Circular trajectory; (d) Back-and-forth trajectory.

The algorithm applied in our system is depicted in Figure 12.

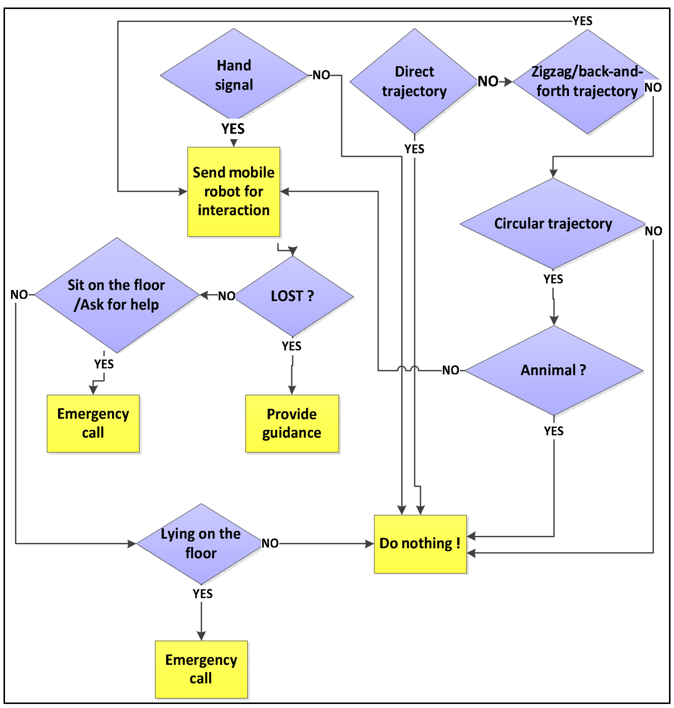


Figure 12. Illustration of the system's algorithm.

A. *Detection of a Direct Trajectory*

If the FID detects a direct trajectory T_d, TAID will orient itself towards the subject in movement without approaching the same.

The rotation angle α(TAID, moving subject) is calculated subtracting the β(TAID, X axis) angle and the γ(moving subject, X axis) angle.

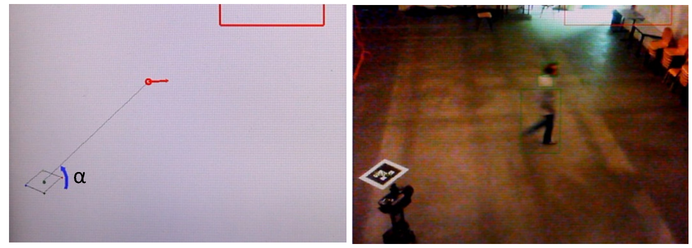


Figure 13. Illustration of the orientation of the robot TAID.

In Figure 13, the orientation of the mobile agent is illustrated. The red arrow represents the motion direction of the subject in movement, which allows knowing if the subject is directing itself towards an area of danger. The blue arrow represents the rotation that the robot has to execute, in order to observe the subject in movement.

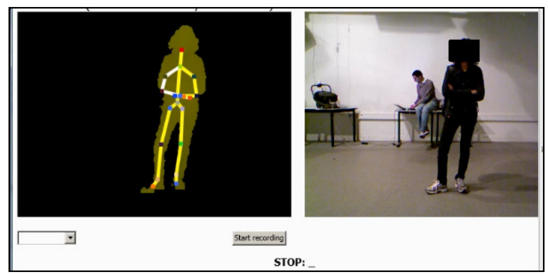


Figure 14. Illustration of TAID verifying the hand sign and the body pose.

TAID is always checking for hand signals asking for help, as it is shown in Figure 14.
However, the robot may detect a falling subject pointed by the FID, employing the 3D camera *Kinect*^{TM}.

B. *Detection of a Nondirect Trajectory*

If the fixed camera detects a zigzag trajectory T_z, the robot will receive the information of intervention with the detected individual. The mobile agent TAID will interfere and progress towards the subject, in order to verify its condition and to interact with the same. After having moved, TAID will collect the following information: size, corpulence and pose of the subject (standing, sitting on the ground or lying down).

In the case of a standing subject, the robot TAID will only ask if the subject needs help and if yes, what kind of help. The interaction is done on the level of the tactile screen of the robot. If a subject is sitting on the ground, TAID will transfer the information to the FID. The agent TAID will send an emergency alert, in order to get a human intervention, when a subject is lying in the ground.

V. RESULTS

In our experiments, each trajectory was tested 9 times, which allowed obtaining 91.67 % successful results. Examples of trajectory detection are given in Figure 15.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

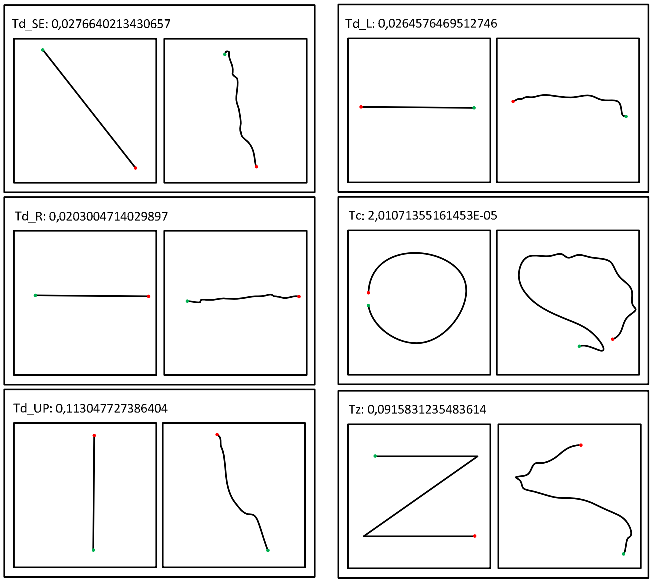


Figure 15. Examples of trajectory detection.

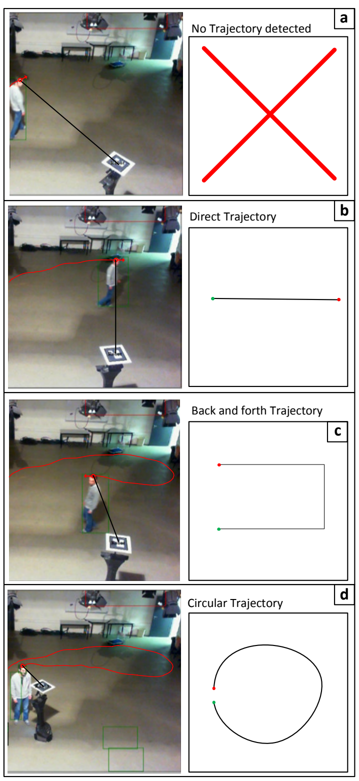


Figure 16. Illustration of trajectory detection and interaction of TAID with a person.

In Figure 16-a, the FID is not detecting any trajectories because of insufficient points.

In Figure 16-b, the FID is detecting a direct trajectory. Hence, the robot is waiting, in case the subject fells or makes a hand sign.

In Figure 16-c, the FID has detected a back-and-forth trajectory. Subsequently, the robot TAID approaches the subject in movement.

In Figure 16-d, the FID has detected a circular trajectory (same result if zigzag trajectory is detected). Consequently, the robot TAID is sent to interact with the subject, asking if help is needed.

If the object in movement is an animal, the intervention will be cancelled, as TAID is able to distinguish between humans and animals.

VI. Conclusion and Future Work

The results show that assistance on a subject by human intervention is only needed in the most critical cases. The analysis of trajectory behaviour helps to make a distinction between decided and hesitating subjects. This pre-analysis allows the mobile agent to approach the hesitating subject, in order to interact as fast as possible. Therefore, the definition of a trajectory can be changed, according to the environment. Having two different views helps the system to confirm unclear cases, such as detected animals or vehicles.

The future work will consist in testing our system with more than one subject, as well as with more agents in the field. This will allow the development of the communication between the agents and the determination of how interaction priorities can be managed between the present agents.

References

[1] R. Rosales, and S. Sclaroff. Trajectory Guided Tracking and Recognition of Actions. BU-CS-TR-99-002: Publisher Boston University Boston, MA, USA ©1999, 1999. http://dcommon.bu.edu/xmlui/bitstream/handle/2144/1779/1999-002-trajector-guided-tracking-and-action-recognition.pdf [retrieved: April, 2012].

[2] R. Bodor, B. Jackson, and N. Papanikolopoulos.Vision-Based Human Tracking and Activity, in Proc. of the 11th Mediterranean Conf. on Control and Automation, Vol. 1 (2003). http://mha.cs.umn.edu/Papers/Vision_Tracking_Recognition.pdf [retrieved: April, 2012].

[3] R. Annicchiarico, C. Barrué, T. Benedico, F. Campana, U. Cortés, and A. Martínez-Velasco, The i-walker: an intelligent pedestrian mobility aid, in ECAI 2008 - 18th European Conference on Artificial Intelligence, Patras, Greece, July 21-25, 2008, Proceedings, pp. 708–712, 2008. ISBN: 978-1-58603-891-5. http://www.diagnostic-walker.es/pubs/iWalkerPatras.pdf [retrieved: April, 2012].

[4] K. Nagai, I. Nakanishi, and H. Hanafusa, Assistance of self-transfer of patients using a power-assisting device, in IEEE Int. Conference on Robotics and Automation, p. 4008 U4015, 2003. ISBN: 0-7803-7736-2.

[5] D. Chugo, T. Asawa, T. Kitamura, S. Jia, and K. Takase, A motion control of a robotic walker for continuous assistance during standing, walking and seating operation, in IROS, pp. 4487–4492, 2009. ISBN: 978-1-4244-3803-7. http://cdn.intechopen.com/pdfs/8624/InTech-A_motion_control_of_a_robotic_walker_for_continuous_ass

istance_during_standing_walking_and_seating_operation.pdf [retrieved: April, 2012].

[6] P. Méderic, V. Pasqui, F. Plumet, P. Bidaud, and J. Guinot, Design of a walking-aid and sit-to-stand transfer assisting device for elderly people, in 7th Int. Conference on Climbing on Walking Robots (CLAWARŠ04), 2004. http://www.isir.upmc.fr/files/2004ACTI97.pdf [retrieved: April, 2012].

[7] C. Riman, E. Monacelli, I. Mougharbel, A. El Aij, Y. Alayli, "A Multi-Interface Platform System for Assistance and Evaluation of Disabled People", First Number Special Issue on Assistive Robotics, Applied Bionics and Biomechanics Journal, vol. 8, pp. 55-66, 2010.

[8] E. Monacelli, F. Dupin, C. Dumas, P. Wagstaff, "A review of the current situation and some future developments to aid disabled and senior drivers in France", Elsevier Masson BioMedical Engineering and Research, IRBM, 30, pp 234–239, 2009.

[9] Microsoft, *Xbox 360 Kinect Sensor Manual*, October 2010. http://www.xbox.com [retrieved: April, 2012].

[10] C. Schüldt, I. Laptev, and B. Caputo, Recognizing Human Actions: A Local SVM Approach, 17th International Conference on Pattern Recognition (ICPR'04) - Volume 3, 2004. ftp://ftp.nada.kth.se/CVAP/users/laptev/icpr04actions.pdf [retrieved: April, 2012].

[11] M. Khalili, E. Monacelli, and Y. Hirata, Identification of pedestrian characteristics for assistive systems, in *IEEE/SICE International Symposium on System Integration (SII 2011)*, 2011.

Comparative Analysis of Wireless Devices

Maria Astrakhantceva
Samara State Aerospace University, Moskovskoe sh., 34, Samara, 443086, Russia
E-mail: astrakhantseva.ma@megafonvolga.ru

Andrei Sukhov
Samara State Aerospace University, Moskovskoe sh., 34, Samara, 443086, Russia
E-mail: amskh@yandex.ru

***Abstract*—In this paper, we attempt to compare the quality of communication wireless equipment manufactured by different vendors. A range of smart devices operating in the various standards are available on the market today. Special attention is paid to the *WiFi* standard as the most popular and affordable due to cheap hardware. The dependence of packet loss on the signal power level and load bus is described using linear approximation. The resulting model coefficients have been found in the experiment and compared. A description is given of the experiment, which was conducted for several types of equipment, and the main features of the equipment are identified.**

Keywords-quantitative comparison of the WiFi equipments; packet loss; connection quality of the wireless standard IEEE802.11b/g/n

I. INTRODUCTION

Modern mobility requires fast connection to the global Internet network at any point in the shortest time. The most popular and cheapest type of wireless connection is a wireless Ethernet of Wi-Fi standard (*IEEE 802.11g/n*) [6]. A Wi-Fi network can be easily deployed in any building, including historic buildings where cabling is not possible.

The question of choosing the best wireless device for each user arises during the connection [12], [20]. Existing methods of comparison are based on qualitative methods [4], [22], which rely on subjective human experience. The purpose of this work is the construction of a universal analytical model to allow us to compare quantitatively several parameters that describe objectively the quality of wireless connection.

For example, streaming has been hard to implement in wireless networks using the old standards (*Wi-Fi, GSM, 3G*) because of a large percentage of packet loss. The percentage of packet loss should not exceed 0.5% for the quality of the wireless connection to be good. Streaming is generally not possible if the percentage of packet loss is over 1.5% [15]. Therefore, in this paper special attention will be given to minimizing the percentage of packet loss.

This paper is organized as follows. First, in Section II, we note the basic parameters affecting the quality of network connectivity, and based on these, we construct a simple analytical model that allows a quantitative comparison of several parameters describing objectively the quality of the wireless connection. In order to verify the model, an experiment in the wireless network is proposed in Section III. The kernel of this experiment will be the investigated wireless equipment. Finally, Section IV discusses the ramifications of the experiments.

II. THE MODEL OF COMPARATIVE ANALYSIS

According to the IETF standard from RFC 2544 [1], the quality characteristics of a TCP/IP based network are described by the following parameters [5]:

- D - packet delay (measured in milliseconds, *ms*);
- p - packet loss (measured in percent, %);
- j - delay variation or network jitter (measured in milliseconds, *ms*);
- B - available bandwidth for end to end connections (measured in Megabits per second, *Mbps*).

For wireless networks the main parameter of network connection quality is packet loss. In [13], it was shown that video quality in wireless networks due to packet loss of 80% and only 20% of network jitter. The aim of this work is to identify the main parameters of wireless networks that have the greatest impact on the quality of network connection (packet loss).

Preliminary tests in wireless networks show that the following parameters have the greatest impact on the quality of wireless network connections:

- I is the signal power of the wireless network, measured in *dBm* where the zero reference point corresponds to one milliwatt;
- B is the average load of the wireless switch.

Each type of wireless network equipment can communicate within certain limits. The range of values for signal strength and network load will form an operations region (gray area in Fig. 1)

$$S \in \left[\overline{I_{min} : I_{max}}; \overline{0 : B_{max}}\right], \qquad (1)$$

where

- I_{max} is the maximum signal power, which can be obtained by the client;
- I_{min} is the minimum power level at which the network connection will still be carried out;
- B_{max} is the maximum load of the wireless hub, that is, the total rate of all channels passing through a wireless device.

This area is an inherent characteristic of wireless equipment and should be displayed in its documentation. The limit values of the operation region are determined by packet loss rate p,

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

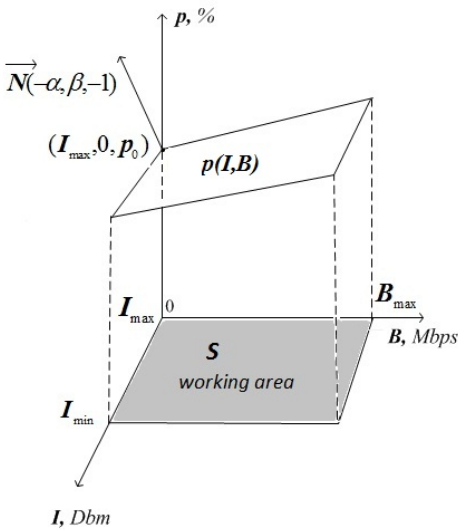


Fig. 1. The working plane $p(I, B)$ and operations region S

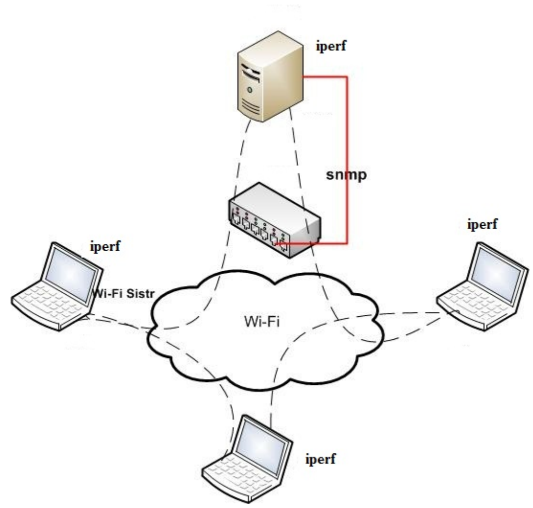


Fig. 2. Scheme of experiment with *iperf* utility

which should not exceed 15% [11]. For this level of error, it is easy to find the upper limits of the operation region, respectively, I_{min} and B_{max}. Averaged values of packet loss $p(I, B)$, depending on the channel loading B and power of wireless signal I, will form some kind of surface (I, B, p), which characterizes the quality of the network connection.

In the first approximation, dependence $p(I, B)$ can be described by a linear function. In other words, a surface (I, B, p) is a plane which can be described by a normal vector $\vec{N}$ and the initial point $(p_0, I_{max}, 0)$. This surface is shown in Fig. 1.

Thus, we obtain a set of parameters that will be evaluated as a wireless connection:

- I_{max} is the maximum signal power;
- I_{min} is the minimum signal power;
- B_{max} is the total loading of the wireless equipment;
- $\vec{N}\,(-\alpha, \beta, -1)$ is the normal vector to the working plane;
- $(I_{max}, 0, p_0)$ is the starting point.

The equation of the working plane can then be written as:

$$\vec{N}\left(\vec{R} - \vec{R_0}\right) = 0, \tag{2}$$

where

- $\vec{R}\,(I, B, p)$ is a vector that characterizes the wireless network state at the current time;
- $\vec{R_0}\,(I_{max}, 0, p_0)$ is a vector characterizing the wireless network state at maximal signal power.

Turning to the variables, we obtain a linear dependence for function $p(I, B)$:

$$p = p_0 + \alpha\,(I_{max} - I) + \beta B \tag{3}$$

where, coefficients α and β reflect the linear relationship between packet loss p *vs* signal strength I and packet loss p *vs* the bus load of switch B correspondingly.

The question arises, in which case a linear approximation can be used. During the experiment three values I_i, B_i, p_i are measured. Next, we try to formulate the conditions that should be applied to experimental data in order to be able to use a linear approximation of Eqn. (3). As main criterion, it is advisable to choose the following condition: the maximum change of value p on the working area *S* of Eqn. (1) should be two times higher than the experimental error. This condition in an analytical form can be written as follows:

$$\max \Delta p^{theor} > 2\Delta p^{exp}. \tag{4}$$

The values of the linear coefficients α and β of Eqn. (3) are found by the least squares method from the minimum to the experimental error:

$$\Delta p^{exp} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - \alpha I_i - \beta B_i)^2}, \tag{5}$$

where α and β should satisfy the condition

$$\sum_{i=1}^{N} p_i = \alpha \sum_{i=1}^{N} I_i + \beta \sum_{i=1}^{N} B_i, \tag{6}$$

and they should be run through a range of possible values with a step $\Delta\alpha$ and $\Delta\beta$, that is easy to implement for a grid with parallel programming. Then the condition of sufficiency of the linear approximation looks like:

$$|\alpha(I_{max} - I_{min})| + |\beta(B_{max} - B_{min})| \geq 2\sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - \alpha I_i - \beta B_i)^2}, \tag{7}$$

N is the number of measurements.

III. Experiment Tests

In order to verify the above model and to calculate the model coefficients α and β, it is necessary to organize an experiment with the wireless network, the core of which will be the investigated equipment. The loading of wireless

Copyright (c) IARIA, 2012.
ISBN: 978-1-61208-225-7

switch B is removed using **snmp** protocol (simple network management protocol [3]) directly with the wireless devices. The scheme of the experiment is shown in Fig. 2.

Wireless switch is connected to the network via a twisted pair. During the experiment, we study only one wireless communication channel between the switch and the laptop. In the proposed scheme, only the switch is replaced under test, all tests were performed under equal conditions. In this experiment, we compare the packet loss at equal signal power and switch load.

In order to measure the signal power of the wireless networks, the following utilities were installed on the computer and tested: *NetStumbler*, *inSSIDer*, *Wi-Fi SiStr*, *Wi-Fi Hopper*. Next, we present a comparative analysis of these products:

- **NetStumbler** is a tool for Windows that facilitates detection of Wireless LANs using the *802.11b*, *802.11a* and *802.11g* WLAN standards. It runs on Microsoft Windows operating systems from Windows 2000 to Windows XP. The program is commonly used for: wardriving, verifying network configurations, finding locations with poor coverage in a WLAN, detecting causes of wireless interference, detecting unauthorized (rogue) access points, aiming directional antennas for long-haul WLAN links [18].
- **inSSIDer** like *NetStumbler* uses active scanning techniques, and all found the information on access points shown in the table [7].
- **Wi-Fi SiStr** works steadily and is compatible with other software. It is particularly suitable for our experiments [17].
- **Wi-Fi Hopper** can display details like SSID, network mode, encryption type, RSSI, frequency and channel, amongst numerous others, for a complete picture of the environment. It is easy to filter out classes of networks by using the network filters. Additionally, a GPS device can be used for reviewing the approximate locations of the detected access points [9].

From all the programs listed above, only *Wi-Fi SiStr* works stably in conjunction with the software (*Videolan* [21], *Wireshark* [16]). Other programs use active monitoring of the network and do not allow parallel operation of other applications, namely, to invalidate a connection to a wireless network.

Wireshark software was used for analysis of packet loss. This allows the user to view all the network traffic in a real-time regime. *Wireshark* distinguishes between the structures of different network protocols, and therefore allows us to parse the network packet, showing the value of each protocol field at any level. For our experiments, the use of a built-in *Wireshark* RTP flow analyzer was important to show the percentage of packet loss.

On a laptop connected to a wireless network, two programs were installed for network monitoring:

- **iperf** for information about the quality of the connection (j, p, B) [8],
- **Wi-Fi SiStr** for measuring the signal strength (I),
- **Wireshark** for measuring the percentage of packet loss,
- **Videolan** for organizing the RTP stream.

With the utility *iperf* installed on the server we are able to create the required network loading with different hosts. On the server side, the *iperf* utility was run with the following options:

iperf -s -u -i.

On the laptop side (client), it starts with the following options:

iperf -c node2 -u -b 2m,

where *node2* is IP address of the server, *m* is the required switch loading. The scheme was originally used by testing *iperf*, as shown in Fig. 2. In order to change the signal levels, the distance between the laptop and switch is changed. In the experiment, we investigated wireless switches produced by *3COM (Model 7760)* and *D-Link (DAP-1150)*.

Because we were limited in time, only two equipment types of two different manufacturers, *D-Link* and *3 COM*, have been tested. In the near future different wireless equipment manufactured *Cisco* and *Juniper* is planned for testing.

A further feature of the utility *iperf* became clear during the experiment. During measurement, it seeks to minimize the packet loss by reducing the connection speed. Therefore, the characteristics of the equipment are the dependence of the data transfer rate of signal power and given value B. The experimental setup was upgraded to detect packet loss (see Fig. 3).

The measurements were performed with RTP/UDP streams [19].For our purposes, it is sufficient to measure only the percentage of lost packets in the stream. Since inside the RTP/UDP stream packets are numbered, then packet loss is easy to fix at constant speed of transmission. It should be noted that this study does not address the quality of the percept video [10], [2], as well as different types of encryption.

Video was selected and encoded for transmission on the network at different speeds (500, 1000, 2000, 5000, 10 000 *Kbps*). Streaming was implemented using the *Videolan* server (vlc) and the receiving stream data was recorded using the network analyzer *Wireshark*. The RTP streams analyzer integrated into this software package shows automatically the percentage of packet loss, which varies depending on the signal I.

IV. THE RESULTS OF MEASUREMENTS

For RTP streams we obtained three values p, I, B, where I is measured by *Wi-Fi SiStr*, B is wondered how the flow rate of during video. The data were processed in accordance with the algorithms given in Section II, and the data are summarized in Table I.

Parameters α and β were calculated as a result of testing with different loads of equipment and signal levels. They characterize the quality of wireless equipment and satisfy Eqns. (5) and (6).

For a wireless switch produced by *3COM (model 7760)*, measuring the level of the signal network shows that the maximum power I_{max} does not exceed -25 *dBm*. When the

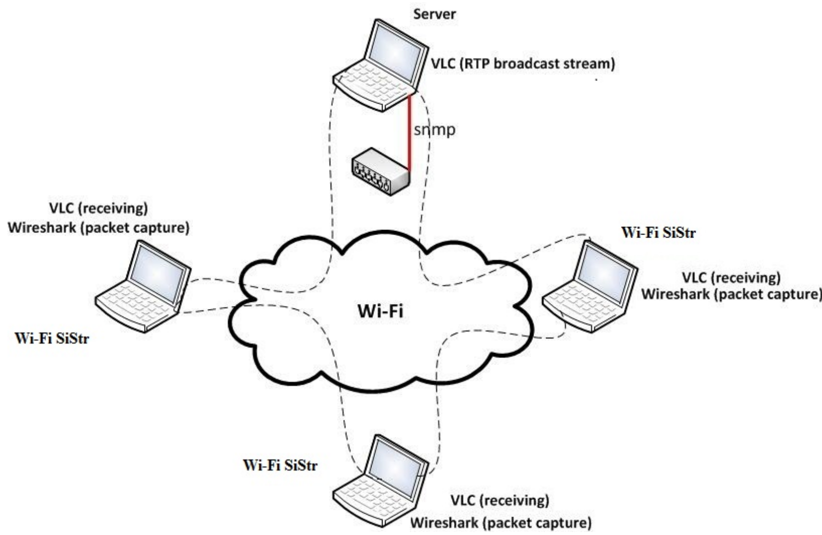


Fig. 3. Scheme of experiment with RTP/UDP stream

TABLE I
BASIC PARAMETERS OF ANALYTIC MODEL

Switch	I_{max}, dBm	I_{min}, dBm	B_{max}, Mbps	α, dBm^{-1}	β, Kbps^{-1}
3COM	-30	-80	25	$(1.0 \pm 0.3) \times 10^{-3}$	$(1.0 \pm 0.6) \times 10^{-6}$
D-Link	-15	-70	25	$(2.2 \pm 1.2) \times 10^{-3}$	$(2.0 \pm 1.6) \times 10^{-6}$

signal level I_{min} = -80 *dBm*, connection terminates. Maximum load B_{max} does not exceed 25 *Mbps*. The quality of communication depends largely on the power of the received signal and not on the speed of RTP streams.

The *D-Link* wireless switch *(model DAP-1150)* shows greater signal power I_{max} = -15 than the *3COM*. Other values do not differ greatly from the *3COM* switch. Measurements for the wireless routers produced by *D-Link* and *3COM* are given in Table I.

Values of α and β coefficients represent the slope angle of the line in the planes (p, I) and (p, B). Comparison of α and β coefficients on the working area indicates that the signal power defines the communication quality by 80 percent. In general, the equipment *3COM* showed that its performance in more than two times better than the competitor *D-Link*.

It should be noted that the network configuration of *IEEE 802.11n* is optimized to reduce the percentage of packet loss. When reducing the power of the received signal, the baud rate automatically drops to a value at which packets are no longer lost. The real network load is reduced by 3-4 times compared with that given by the utility *iperf* option *m*. Thus, the communication quality at comparable settings when using the *D-Link* router is almost an order of magnitude lower.

The TCP/IP connection is considered to be good if packet loss does not exceed 0.5% [2], [14]. The obtained result is at least an order lower than the good level.

In operating the model, experience of networks of standard *IEEE 802.11g* was used for the network connections, which are characterized by a significant percentage of packet loss. The *IEEE 802.11n* standard, however, is characterized by a small percentage of errors, so the model presented here does not describe the quality of the network connection well and needs modification.

The experimental setup is shown in Fig. 3, during the experiment to evaluate the bandwidth used by the utility *iperf*. The essence of the experiment is that for a given bandwidth is a real load is considerably less, depending on the capacity of the wireless signal. Experiments have shown that much better equipment can be described by the following relationship:

- B_{real} is a real value bandwidth using the utility *iperf*;
- B_{iperf} is a value indicating the key m in the utility *iperf*;
- I *Wi-Fi Sistr* is received signal strength.

The experimental results are presented in graphical form depending on the achieved bandwidth of the signal power are shown in Fig. 4. For example, when using *iperf* with bandwidth B_{iperf}=20 *Mbps* and signal level I = -70 *Dbm*, the real speed of data transmission in a wireless network will be

- for **3COM** B_{real}^{3COM} = 6,8 *Mbps*

and

- for **D-Link** B_{real}^{D-Link} = 800 *Kbps* (see Fig. 4).

From this we can conclude that the quality of communication at comparable parameters when using a *D-Link* router is of a lower order.

V. Conclusion

In this paper, a simple analytical model is constructed to compare the quality of wireless networks. Several parameters are selected for quantitative comparison of the investigated

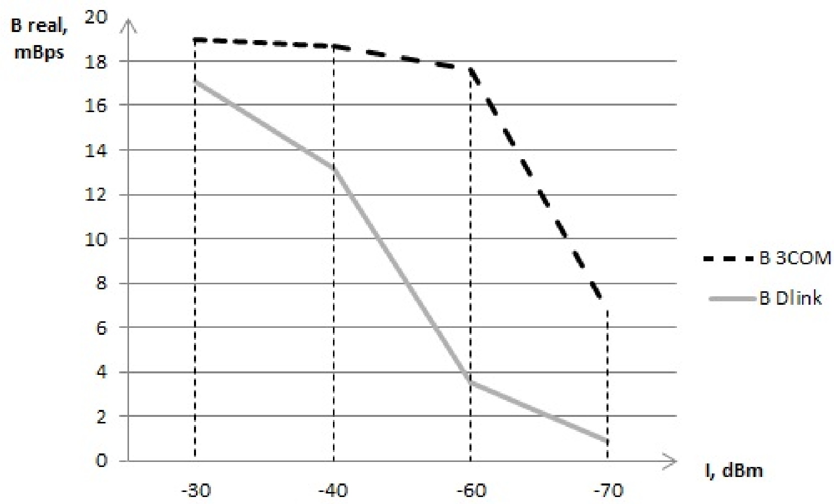


Fig. 4. Real iperf bandwidth

equipment. These options allow us to objectively describe the quality of the wireless standard *IEEE802.11b/g/n*. The experiments have confirmed the suitability of the model for assessing the quality of wireless equipment. Also, the study revealed the superiority of the *3COM* wireless device over the *D-Link* equipment. However, our model needs to be clarified, as *IEEE 802.11n* networks are characterized by a low percentage of losses that can be achieved by reducing the data rate.

REFERENCES

[1] S. Bradner and J. McQuaid, *Benchmarking Methodology for Network Interconnect Devices*, RFC 2544, March, 1999.

[2] P. Calyam, M. Sridharan, W. Mandrawa, P. Schopis, *Performance Measurement and Analysis of H.323 Traffic*, Passive and Active Measurement Workshop (PAM), Volume 3015, April, 2004, pp. 137-146.

[3] J.D. Case, M. Fedor, M.L. Schoffstall and J. Davin, *A Simple Network Management Protocol (SNMP)*, RFC 1157, May, 1990.

[4] P. Deshpande, X. Hou, and Samir R. Das, *Performance Comparison of 3G and Metro-Scale WiFi for Vehicular Network Access*, in Proc. of Conference on Internet Measurement, 2010, pp. 301-307.

[5] M. Jain and C. Dovrolis, *End-to-end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput*, IEEE/ACM Transactions on Networking 11(4), August, 2003, pp. 537-549.

[6] IEEE 802.11 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications (Part 11), 1999.

[7] inSSIDer, http://www.metageek.net/products/inssider.

[8] Iperf, http://sourceforge.net/projects/iperf.

[9] WiFi Hopper, http://wifihopper.com.

[10] J. Klaue, B. Rathke, and A. Wolisz,*"Evalvid-a framework for video transmission and quality evaluation,"*in Proc. 13th Int. Conf. Modeling Techniques and Tools for Computer Performance Evaluation, Urbana, IL, Volume 2794, September, 2003, pp. 255-272.

[11] A. Kostuch, K. Gierlowski and J. Wozniak, *Performance analysis of multicast video streaming in IEEE 802.11 b/g/n testbed environment*, in Proc. IFIP of Conference on Wireless and Mobile Networking, Volume 308, September, 2009, pp. 92-105.

[12] A. Lyakhov and V. Vishnevsky, *Comparative Study of 802.11 DCF and its modification in the presence of noise*, Wireless Networks 11, 2005, pp. 729-740.

[13] E. Sagatov, A. Sukhov and P. Calyam, *Influence of Distortions of Key Frames on Video Transfer in Wireless Networks*, in IEEE International Symposium on Image/Video Communications over Fixed and Mobile Networks (ISIVC), May, 2010.

[14] Leandro Melo de Sales, Hyggo O. Almeida and Angelo Perkusich, *On the Performance of TCP, UDP and DCCP over 802.11g Networks*, in Proceedings of the 2008 ACM symposium on Applied computing, 2008, pp. 2074-2078.

[15] J. Seyba, B. Mullins and G. Bonafede, *Audio-video capacity of an IEEE 802.11g wireless LAN*, 2007 International Symposium on Collaborative Technologies and Systems (CTS 2007), Orlando FL, May, 2007, pp. 372-378.

[16] Wireshark Foundation: Wireshark. Go Deep, http://www.wireshark.org/.

[17] Wi-Fi SiStr, http://www.softpedia.com/get/Multimedia/ Audio/Other-AUDIO-Tools/WiFi-SiStr.shtml.

[18] NetStumbler, http://www.netstumbler.com/.

[19] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, RTP: A Transport Protocol for Real-Time Applications, RFC-1889, 1996.

[20] A. Symington, *A Hardware Testbed for Measuring IEEE 802.11g DCF Performance*, Modeling, Analysis and Simulation of Computer and Telecommunication Systems, 2009, MASCOTS '09, pp. 61-67.

[21] VideoLAN Team: VideoLAN, Free Streaming and Multimedia Solutions for All OS!,http://www.videolan.org/.

[22] Paul A Zandbergen, *Accuracy of iPhone Locations: A Comparison of Assisted GPS, WiFi and Cellular Positioning*, Transactions in GIS, 13, 2009, pp. 5-26.

TRIPZOOM: A System to Motivate Sustainable Urban Mobility

Paul Holleis, Marko Luther, Gregor Broll, Hu Cao
DOCOMO Euro-Labs
Munich, Germany
{luther, holleis, broll, h.cao}
@docomolab-euro.com

Johan Koolwaaij, Arjan Peddemors, Peter Ebben, Martin Wibbels
Novay
Enschede, The Netherlands
{johan.koolwaaij, arjan.peddemors, peter.ebben, martin.wibbels}@novay.nl

Koen Jacobs, Sebastiaan Raaphorst
Locatienet
Amsterdam, The Netherlands
{koen.jacobs, sebastiaan}
@locatienet.nl

***Abstract*—The accelerating growth of cities provides new challenges for urban planning, especially related to transportation. Many existing infrastructures already operate at their limits during rush hours or large events. They need to be used more efficiently as they often cannot be extended without considerable negative effects for citizens such as increasing pollution or travel costs. We describe a framework that enhances mobility data from existing urban infrastructure with data from participatory sensing. Based on derived mobility patterns and direct feedback via social networks, a dynamic incentive system is provided to positively influence mobile behavior on a personal and city-wide scale beyond regular urban planning. We present the design and architecture of such a system and point out critical issues and proposed solutions in an initial implementation aimed towards real world tests in several cities.**

Keywords - sustainable traffic; urban mobility; personal mobile sensing; incentives; social networks.

I. INTRODUCTION

Today, more than 50% of the world's population lives in cities and the UN forecasts a further increase to 70% by 2050 [1]. This *metropolisation* creates many challenges for the world's cities [2]. Current transport infrastructures often cannot cope with the load anymore especially during peak times. They have to be used more sustainably as extensions incur considerable negative effects for citizens such as increasing pollution and direct or indirect travel costs.

The European FP7 project "Sustainable Social Network Services for Transport" (SUNSET [3]) is developing and evaluating the TRIPZOOM system [4] that implements a new approach to urban mobility management. The assumption is that sharing personal mobility patterns via social networks and the use of incentives can encourage citizens to utilize sustainable forms of transportation and to generate a win-win situation for all involved stakeholders (Fig. 1). Citizens can optimize their mobility needs using recommendation and personalized traffic services from the city authority, share travel related information with buddies on social networks (like the community-based traffic app WAZE [5]) and get rewarded for sustainable behavior. The city authority receives detailed mobility profiles of citizens, relevant to the assessment of current infrastructure use and future mobility needs, and achieves optimizations by offering incentives. Communities, like employees of a company or participants of a car sharing system, are supported in networking and in enlarging their groups. Third party service providers can tap into the data to create novel offerings and integrate with others through common incentive structures (like the location-based incentive application FOURSQUARE [6]).

To reach the project goals, we are realizing a technical and socially supportive environment called TRIPZOOM that complements urban infrastructure sensing with participatory mobility sensing. TRIPZOOM is based on our competences gained by developing and operating the context-aware community application IYOUIT [7], the mobility capturing application TRAVELWATCHER [8], as well as the traffic jam alerting application FILEWEKKER [9]. In contrast to other real-time mobility monitoring systems focusing on anonymous, aggregated data (such as [10]), TRIPZOOM generates detailed personal mobility patterns. These allow users to zoom in on trip details and raise awareness by computing trip overviews and eco feedback [11] that can be shared with friends or the community.

It is often assumed that direct feedback can positively influence the mobility behavior of individuals and that social influence is probably the most powerful factor behind human behavior change. We will use the TRIPZOOM system to evaluate these assumptions and to develop feasible and successful incentives by operating living labs in three European cities.

In the following, Section II describes the incentive and social networking aspect on which TRIPZOOM is built. Section III highlights important aspects of the mobile sensing part while Sections IV and IV provide details about the architecture and implementation. We conclude with a description of the current status and next steps within the project.

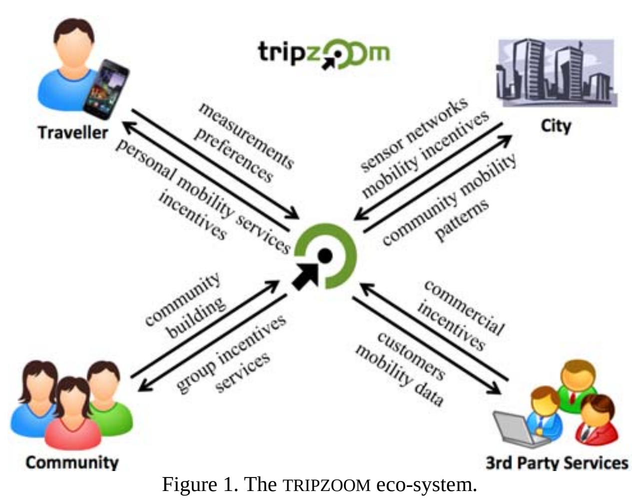


Figure 1. The TRIPZOOM eco-system.

II. INCENTIVES AND SOCIAL NETWORKING

TRIPZOOM aims at providing different types of incentives to relevant target groups at relevant places and in relevant situations. To incorporate the most effective types of incentives, we conducted an analysis of individual travel behavior [12]. This research implies that influential incentives should be based on the following areas:

- *Time:* efficient use, control, saving, and planning
- *Money*: save (e.g., discounts on transportation tickets) or even generate (e.g., coupons)
- *Information*: receive (real-time, personalized) information about progress, travel alternatives, …
- *(Social) recognition*: of being green and healthy/fit; give feedback to and receive feedback from others

In TRIPZOOM, incentives appear mostly in form of a challenge defined by a set of rules that users have to fulfill and a reward that they can earn (e.g., take the bike instead of the car to get a coupon financed by a bicycle store). To be able to evaluate the effect of different (combinations of) incentives, TRIPZOOM offers dynamic and controllable incentive management. It allows for generating and placing incentives in real-time during the entire runtime of the system. This means that weak or unintended side effects of incentive offers [13] can be detected early and that incentive operators such as city representatives can alter incentives accordingly.

For this purpose, a component called city dashboard (see Section IV) is being developed that will offer a live view on various aggregated statistics and anonymized data of the users within a city. It can be used to efficiently check and evaluate the effect of incentive measures. This approach can be used to change overall mobility behavior (e.g., fewer cars in the city) or to target specific goals, e.g., optimizing the use of alternative travel modalities before and after large events.

TRIPZOOM uses a point scheme as one way to integrate various offers, and to open the incentive system to third parties: each incentive is worth a certain number of points, which can be exchanged for certain rewards, for example information or monetary items such as access to fare reductions. This point scheme also makes it easier for users to track their progress as well as to share and compare it with friends and colleagues (Fig. 2). This adds competition as a gamification feature [14] and creates a more playful experience. For that purpose, TRIPZOOM offers a set of social networking features that connect to popular sites such as FACEBOOK, FOURSQUARE and TWITTER. That way, existing social networks are used to advertise TRIPZOOM and to allow users to share their progress and achievements with others. These social connections can themselves be seen as a powerful tool for inducing behavior change as people can engage in competitions. Furthermore, all users' privacy and sharing settings are defined with respect to their connections.

III. MOBILITY SENSING

To optimize mobility, one needs to identify individuals that are relevant for the optimization targets under consideration. Therefore, mobility sensing plays a fundamental role in the proposed TRIPZOOM system. To gain profound insight into mobility behavior, TRIPZOOM combines the strengths of two orthogonal approaches: infrastructure sensing that measures vehicle intensities or travel times in urban networks (e.g., roads, rails, bus lanes), and personal mobility sensing that uses mobile phones to track how citizens make use of that network.

Infrastructure sensing includes traffic light and intersection status providing information about traffic intensity, intervals of traffic light phases, and delay times. License plate cameras can estimate dynamic travel times on the roads between any two cameras. Parking areas deliver information about the number of total and available lots. Further data can be real-time public transport information such as the actual location of buses. Infrastructure data is often limited in geographic coverage, and has restricted possibilities to track individuals throughout the whole transportation network.

Personal mobility sensing uses various technologies in citizens' mobile phones (e.g., GPS / Wi-Fi / accelerometer [8]) to detect or deduce the 4 Ws of personal mobility: **w**hen do people move, **w**here (via which route), with **w**hom, and using **w**hich modality (bike, car, etc.). While personal mobility data allows detailed individual tracking, only a small part of the city community can be expected to participate. Still, the goal is to get an overall mobility profile per citizen (Fig. 3) covering a 24/7 period. The use of mobile phone sensors and battery power has to be carefully balanced with the measurement accuracy to be achieved [15].

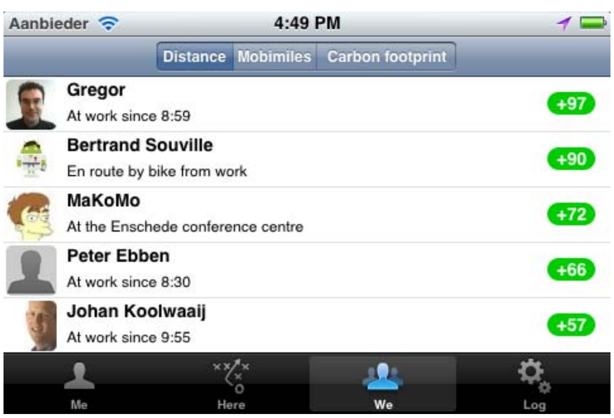


Figure 2. Buddy list including current incentives points.

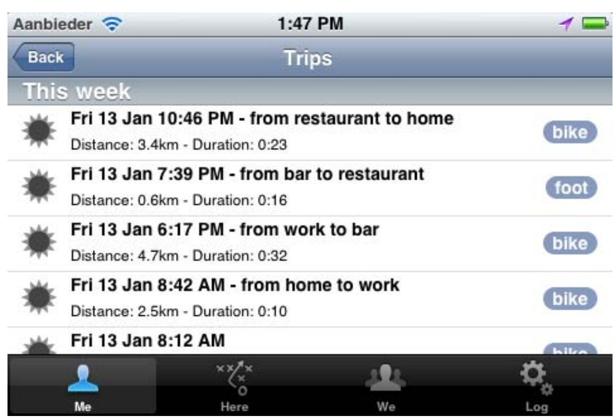


Figure 3. Trips and detected modalities.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

Combining and correlating both approaches relieves these shortcomings by improving the level of detail regarding the geographic and community coverage of the TRIPZOOM measurement approach.

IV. ARCHITECTURE

TRIPZOOM consists of a network of Core Services guarded by Proxy and Security Services. Mobile clients, a web portal, and the city dashboard act as interfaces to the user. Furthermore, a social network infrastructure allows for easy sharing of data (see Fig. 4 and detailed description in [16]).

The service infrastructure delivers incentives (see Section II), mobility monitoring (see Section III), and basic social network services. The security layer allows access to the TRIPZOOM universe only according to user consent.

The mobile application supports sensing of personal mobility by connecting to (sensor) data such as mobile phone location information and offers a means to interact with the user. The Web portal additionally provides visualizations of statistical data as well as comparisons with respect to community mobility patterns. Traffic authorities and researchers use the city dashboard to get an overview of the current mobility situation and have control over applied incentives.

Additionally, TRIPZOOM uses resources offered by external or third party components, e.g., sensors and services to monitor road traffic and to obtain public transportation status information. In the same way, third party applications can request access to services provided by the system components depending on user consent. Six components provide the core service functionality.

Personal Mobility Store (PMS): collects raw measurements from mobile clients and preprocesses them to be input to algorithms such as pattern detection.

Mobility Pattern Detector (MPD): receives data from the PMS and employs sophisticated algorithms to detect patterns for individuals, groups, places, regions, routes, or vehicles such as bus lines or a taxi.

Mobility Pattern Visualizer (MPV): takes the patterns derived by the MPD and turns them into interpretable and easily accessible visualizations depicting long-term statistics, trends and personal, city-wide, or place-oriented performance with respect to set goals.

Relation, Identity, and Privacy Manager (RIP): provides its own social network implementation and organizes the privacy policies of users based on their social relations or ad-hoc groupings computed by the MPD.

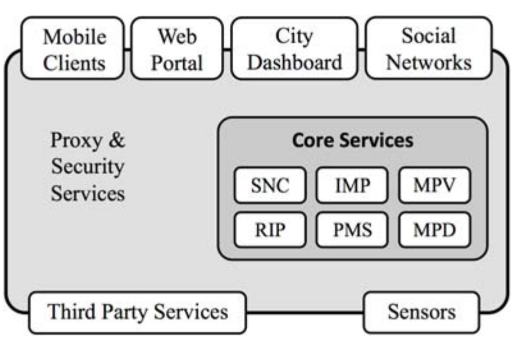


Figure 4. Main system components.

Social Network Connector (SNC): connects the internal social network (RIP) with existing social networks such as FACEBOOK or FOURSQUARE, to facilitate user registration, information sharing (e.g., a notification on the successful completion of an incentive), importing contacts, or showing visualizations from the MPV.

Incentives Market Place (IMP): provides a platform for incentives as rewards, recognition, or real-time feedback to encourage travelers to improve their behavior with respect to the system's and individual's objectives. The IMP matches challenges with mobility patterns from the MPD, profiles and preferences from the RIP, general transport information, and can publish performance and events using the SNC.

V. IMPLEMENTATION

The distributed system is implemented building on Representational State Transfer (REST) and JavaScript Object Notation (JSON). This supports loose-coupling between the components, clear interfaces, and independence of platform and programming language. The security layer uses OAuth [17], a simple mechanism to publish and interact with data that needs access control. It is wide-spread and used by many systems such as FACEBOOK or TWITTER. The TRIPZOOM social network features such as the Social Network Connector build on the open source social networking engine ELGG [18] Using such a platform has the advantages that users are not required to use a specific network and it simplifies the integration with other functionality that the portal and mobile application offer.

A. Mobile Client

Information from mobile device sensors (GPS, Wi-Fi, and GSM measurements) is combined into one model that can deal with the varying accuracy of location measurements. In line with other research in this area (such as [14]), TRIPZOOM uses a smart strategy to record trips in detail but applies less detailed sensing when people are in buildings, or on frequently used routes.

To implement the trip detection in a battery-efficient way, the application listens to 'significant' location changes using Wi-Fi (preferred) or GSM location updates and only enables the battery-hungry GPS if it has reasons to assume that the user has left the (static) location. Attempts to acquire a GPS fix are kept short – if there is no fix within that time, the GPS is turned off and will be re-enabled only after some time. By varying those time intervals, a trade-off is made between consuming battery power and attempting to record the trip as detailed as possible. We refer to [19] for details on this sampling strategy. Our initial tests within a limited user community have shown that battery life is 20 hours on average, with quite some variations related to differences in coverage (GSM, GPS, etc.) and mobile phone use.

The mobile application submits all location recordings to the TRIPZOOM server. All recorded trips are displayed on the mobile application to increase the awareness for personal mobility behavior. Specifically, regular trips, personal places and the mobility footprint of the user are visualized. The latter includes an estimate of the mobility impact in terms of costs, total and lost time, and CO_2 emissions.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

B. Server

On the back-end, the goal is to further enhance the data and to apply smoothing and outlier detection to the location data and to improve trip timings. To gather information on overall modality use and reward calculation, trips need to be stitched and split such that each segment is made of a single modality only. To facilitate modality detection, each route is mapped onto the underlying infrastructure network. Our algorithms estimate the modalities of each trip based on the location measurements as well as derived parameters such as the speed pattern, infrastructure usage, and origin and destination of the trip. The latter are automatically linked to places, to reveal the trip purpose, such as commuting between home and work, and improve the accuracy of modality recognition. Personal places frequently visited by the user in the past, like home or office, are recognized but can be adjusted by the user. Common places, like train station or theater, are mapped to crowd-sourced places as defined, e.g., in FOURSQUARE.

Trip modality recognition applies rules to all gathered information and calculates modality likelihoods. For example, a trip going from one airport to another with an average speed above 150km/h will likely have been made by plane. Users travelling on roads closed for car traffic are likely to walk or cycle depending on their speed pattern, etc. Still, it is important to allow users to manually adapt the recognized modality, in case of any error. The current system has an accuracy of about 77%, which already saves considerable effort compared to manually labeling all trips and is similar to other state-of-the-art approaches [20]. Its algorithm works better on 'walking', 'cycling', 'car', 'plane', and 'train' than on other modalities including 'bus' and 'tram'. An improved approach for better recognizing those modalities is currently in the design phase.

VI. STATUS AND OUTLOOK

SUNSET is an ongoing project. After the identification of the stakeholders, the creation of usage scenarios and requirements analysis [21], the overall system has been designed and the architecture defined as described in the previous section. Several parts of it have already been implemented, connections to external infrastructure and sensor data have been established, and the underlying social network functionality has been provided. Additionally, functional prototypes of the mobile application (for iPhone and Android) are available and a basic web portal is running for evaluation purposes.

The next step after the finalization of the core system development is the start of living labs in three European cities (Enschede, NL, Gothenburg, SE, and Leeds, UK). In cooperation with city councils and public transport data providers, this serves as a test setting and evaluation platform [22] for the defined scenarios, algorithms such as modality and pattern detection, and the power of incentives. To maximize the overall outcome, each living lab will focus on different target groups and follow different evaluation goals [23]. The living labs will be started in a sequence to allow the transfer of the acquired knowledge and experiences from one to the next. Finally, we plan to assess the transferability of the system and living lab results to other cities and environments.

ACKNOWLEDGMENTS

This work has been performed in the framework of the EU FP7-270227 project SUNSET, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues.

REFERENCES

[1] http://esa.un.org/unup/index.asp?panel=1, accessed 5/14/2012

[2] M. Naphade, G. Banavar, C. Harrison, J. Paraszczak, and R. Morris. Smarter cities and their innovation challenges. *Computer*, 44(6): pp. 32–39, June 2011.

[3] http://www.sunset-project.eu/, accessed 5/14/2012

[4] http://www.tripzoom.eu/, accessed 5/14/2012

[5] http://world.waze.com/, accessed 5/14/2012

[6] http//foursquare.com/, accessed 5/14/2012

[7] S. Böhm, J. Koolwaaij, M. Luther, B. Souville, M. Wagner, and M. Wibbels. Introducing IYOUIT. In *Proc. of the Int. Semantic Web Conference (ISWC'08)*, Springer, pp. 804–817, 2008.

[8] W. B. Teeuw, J. Koolwaaij, and A. Peddemors. User behaviour captured by mobile phones. In *Proceedings of the Workshop on Interactive Human Behavior Analysis in Open or Public Spaces (INTERHUB'11)*, 2011.

[9] http://www.filewekker.com/, accessed 5/14/2012

[10] F. Calabrese and C. Ratti. Real time Rome. Networks and Communication Studies. Journal of the IGU's Geography of Information Society Commission, 20(3/4): pp. 247–258, 2006.

[11] J. E. Froehlich, L. Findlater, and J. A. Landay. The design of eco-feedback technology. In *Proceedings of the Int. Conf. on Human factors in computing systems (CHI'10)*, pp. 1999-2008, 2010.

[12] A. Ma, editor. *Feasible and Potentially Successful Incentives*. EU Project SUNSET, Deliverable D3.4, 2012. To Appear.

[13] J. Costa-Font, M. Jofre-Bonet, and S. T. Yen. Not all incentives wash out the warm glow. CESifo Working Paper Series 3527, CESifo Group Munich, 2011.

[14] G. Zichermann, *Gamification by Design*, O'Reilly, 2011.

[15] K. Lin, A. Kansal, D. Lymberopoulos, and F. Zhao. Energy-accuracy trade-off for continuous mobile device location. In *Proc. of the Int. Conf. on Mobile Systems, Applications and Services (MobiSys'10)*, pp. 285–298. ACM, 2010.

[16] S. Poslad, editor. *Service Framework Architecture and Design*. EU Project SUNSET, Deliverable D5.1, 2011.

[17] http://oauth.net, accessed 12/5/2012

[18] http://elgg.org, accessed 12/5/2012

[19] M. Luther, editor. *Initial Mobile Application Design*. EU Project SUNSET, Deliverable D4.1, 2011

[20] V. Manzoni, D. Maniloff, K. Kloeckl, and C. Ratti. Transportation mode identification and real-time CO2 emission estimation using smartphones. Technical report, Massachusetts Institute of Technology, Cambridge, 2010.

[21] S. Poslad, editor. *Preliminary User, System Requirements Review and Specification*. EU Project SUNSET, Deliverable D1.1, 2011.

[22] T. Broens, H. ter Hofte, and P. Ebben. A reference architecture for living lab measurement systems. In *Proc. of the Mobile Living Labs Workshop (MLL'09)*, 2009.

[23] M. Meeuwissen, editor. *Living Lab Plan*. EU Project SUNSET, Deliverable D7.1, 2012. To Appear.

Smart Factory: System Logic of the Project EPIK

EPIK: Efficient Staff Assignment Through Intelligent and Adaptive Cooperation- and Information Management in the Area of Production

Cornelia Hipp, Thomas Sellner, Janina Bierkandt
Fraunhofer Institute for Industrial Engineering
Stuttgart, Germany
cornelia.hipp | thomas.sellner |
janina.bierkandt@iao.fraunhofer.de

Philipp Holtewert
Fraunhofer Institute Manufacturing Engineering and Automation
Stuttgart, Germany
philipp.holtewert@ipa.fraunhofer.de

***Abstract* – The production process in manufacturing industries is getting more complex and demands dynamic solutions, which can react flexible on different events of the production environment (e.g., alerts). Mainly, the static production process could benefit in efficiency, if the product environment would be smart and would delegate recognized tasks intelligently to workers. This idea is the main topic of the research project EPIK. Within this paper, the project is introduced and the main system logic is discussed. A description is included how the context data of the production environment is analyzed, which conclusions are drawn and how a detected task is assigned to the most appropriate worker.**

Keywords: adaptive interfaces; smart production; intelligent environment; system logic; human machine interface; smart factory; information management

I. INTRODUCTION

The research project EPIK (effective staff assignment through intelligent and adaptive cooperation- and information management in the area of production – German original: „Effizienter Personaleinsatz durch intelligentes und adaptives Kooperations- und Informationsmanagement in der Produktion“) improves the production process with help of a smart environment and individual support of workers with help of mobile and adaptive devices.

EPIK collects context data of the production environment, e.g., which dysfunctions occur on which machine and which workers with necessary qualifications are close by. Subsequently, EPIK decides which worker should be ordered to solve the problem and contacts him on his mobile device. In the decision process, EPIK takes into account several aspects, like how much work does the individual worker already has or has the new task high-priority, that the worker should interrupt the current job. When selected, the worker gets individual information to solve the task. This takes place with an adaptive individual mobile device of the worker (compare Figure 1).

The aim of this system is to optimize the production process and to make the staff assignment more efficiently. This will be achieved through three aspects:

Optimal operating grade of resources regarding employees and reduction of non-productive working time: With usage of mobile devices for staff it is possible to make monitoring of the production process without being fixed to a place. A fixed assignment of staff to machine can be canceled (which is currently standard). Therefore, tasks of different machines can be delegated to different staff. Workers will be supported in work with guidance of a mobile device. As a consequence, non-productive time of workers will be reduced.

Accurately fitting of resource management trough intelligent delegation of work tasks to adequate workers: Before delegating the work task, EPIK calculates how much time different workers will approximately need to accomplish the task. Furthermore, the requested qualification will be matched to the deposited qualification of the worker. At last, it will be looked-up, when the worker conducted the same task the last time. EPIK takes these figures into account to calculate an accurate fitting of workers to open tasks.

Raising efficiency or individual job performance with help of individualized and personalized support for the worker:
The mobile device gives context relevant information to solve tasks. The device will display adaptively information to the individual needs.

Thiel et al. [1] point out that a more efficient working place is getting more and more important in western countries, since the demographic change is leading to a lack of qualified staff and that personnel costs have gotten a major amount of the total production cost. Therefore, EPIK addresses this problem thoroughly, since it makes an efficient assignment of tasks to workers and help workers to

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

get continuously qualified through individual support of the mobile device.

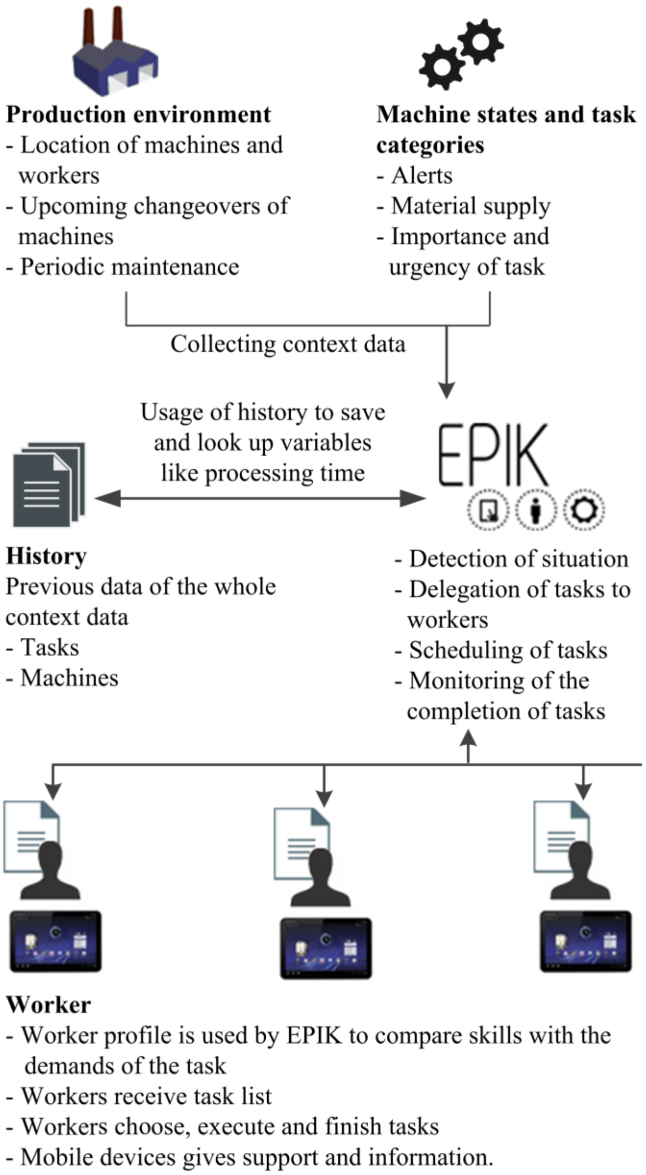


Figure 1. Overview how EPIK captures and analyses the production context with subsequent communication to the worker

II. Situation Detection and Overall EPIK-Algorithm

The aim of EPIK is enhancement of efficiency with help of an optimal operating grade and individual support of the single worker. This demands among other things a dynamic, adaptive user interface with high flexibility. According to Schlegel & Thiel et al. [2] central control systems and static user interfaces have been used for years and the demand for a more flexible possibility in the environment of production is very high. The Federal Government Department of Education and Research in Germany identified this problem as well and claims within a scientific development program [3] that the complexity within production has risen and that there is the necessity for a more flexible production process, which has the ability to observe information, deduce findings from them and to transfer this into change of performance.

EPIK takes this problem into account and collects therefore several context data of the environment, deduce situations from the context data, finds intelligent solutions and contacts the workers flexible on their mobile devices. To find these smart solutions, EPIK chooses the approach of situation detection.

The various context data are changing dynamically and permanently (e.g., the position of the worker) what allegorize high demands at the system. The approach of the situation detection and reaction of EPIK is based on the approach of Kluge [4]:

Scene analysis

Context data is captured continuously and can be from different kind. Context data can be positions (persons, objects, machines, etc.), states (machines, persons, etc.), and task categories. When EPIK detects a situation all context data will be refreshed.

Activity recognition

Subsequently, all captured attributes are put into relationship of each other. For example, moving workers are identified as "walking" due to changing context data. Partial solutions are calculated for the questions, how long does it approximately take each worker to solve the task, which worker is suitable regarding the needed qualification and which worker should practice and learn the task.

Planning

At this step a continuative interpretation of the situation proceeds. Regarding the calculated three lists in the previous step, EPIK makes a weighted comparison to find the best fitting worker to the task (see Section V Assignment of Tasks to Workers). The result of this is going into the process of the scheduling, where EPIK calculates at which position the task should be sorted in the worker-task-list (see chapter VI Scheduling). The scheduling regulates the order of the tasks for the worker and reacts, if the worker refuses the task and start again the selection process to find the next eligible worker. The user gets detailed information on the mobile device about the task.

The context data of EPIK is changing dynamically and is not static (and therefore, not foreseeable). As a result within *scene analysis* and *activity recognition* it is important to work with updated values, to ensure a best possible matching of workers to tasks. This dynamic environment of the system means high demands for the implementation, computing power and algorithms of EPIK. According with Schöning [5] the best solution for dynamic programming is to use a bottom-up-implementation. Schöning [5] writes that partial solutions should be found in advance and saved in a table. The total solution should refer to the partial solutions. This is picked up in EPIK with the previous described lists. The total overview of the overall EPIK algorithm is displayed in the following Figure 2.

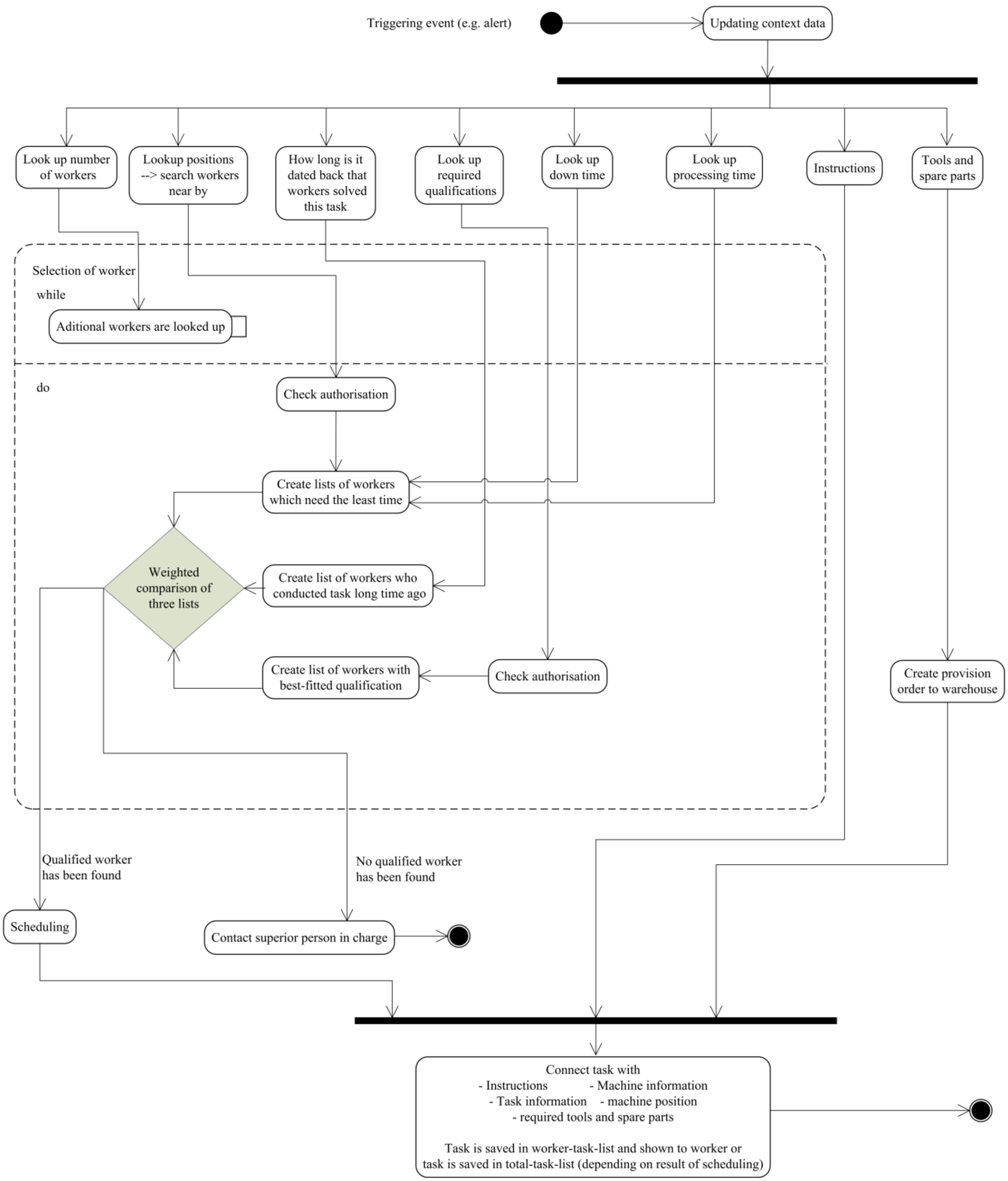


Figure 2. Overview of the overall EPIK algorithm

With the help of the algorithm the goal of reaching an optimal operating grade of resources will be achieved. Workers won't have non-productive working time any more, since the system monitors and detects non-productive working time and will give new tasks to the workers, which are not occupied.

Furthermore, the algorithm ensures an accurately fitting of resource management trough intelligent delegation of

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

work tasks to adequate workers. This is the second approach of the EPIK-project to enhanced efficiency within the industry.

III. System Architecture

To get a better understanding about the system architecture, Figure 3 shows the different system components and their communication ways. The figure demonstrates how the EPIK System is embedded in an IT infrastructure of a production hall and how the System gets the context data, e.g., machine and employee. Especially the machine states are gathered through connecting to the Manufacturing Execution System (MES). Besides the machine conditions the following information are identified as context data. (1) Current values and availability of items in the storehouse are saved in a storehouse items profile. (2) After the situation detection (cf. Section II, Situation detection and overall EPIK-Algorithm) concrete task profiles are generated. (3) At least, every employee is represented by an employee profile that contains several information like the task history, his or her qualification profile, allocated machines, and the current assigned task. To cluster the different responsibilities and tasks of the individual components several managers are implemented. Each of them is modularity implemented so that they can be adapted and extended to other industrial user needs. In the following the certain managers and their tasks are described:

Context Manager: The whole context data that are described above are collected, interpreted and updated by the Context Manager. A further task of the Context Manager is to combine the context data to a concrete situation and chooses the best qualified employee by taking into consideration of the situation parameters. After choosing the best qualified employee the context Manager triggers the Message-Manager that forwards the generated Task to the assigned employee. For that reason this manager plays a key role in the EPIK System.

Recognition Manager: Before the Context Manager interprets the context data, the Recognition Manager monitors the complete machine states by connecting to the MES-Interface. Information are examined and rejected if they are irrelevant for the EPIK-System. In contrast relevant information are gathered and used to generate Task EPIK objects which are forwarded to the Context Manager that completes this object with situation parameters, e.g., from the task profile.

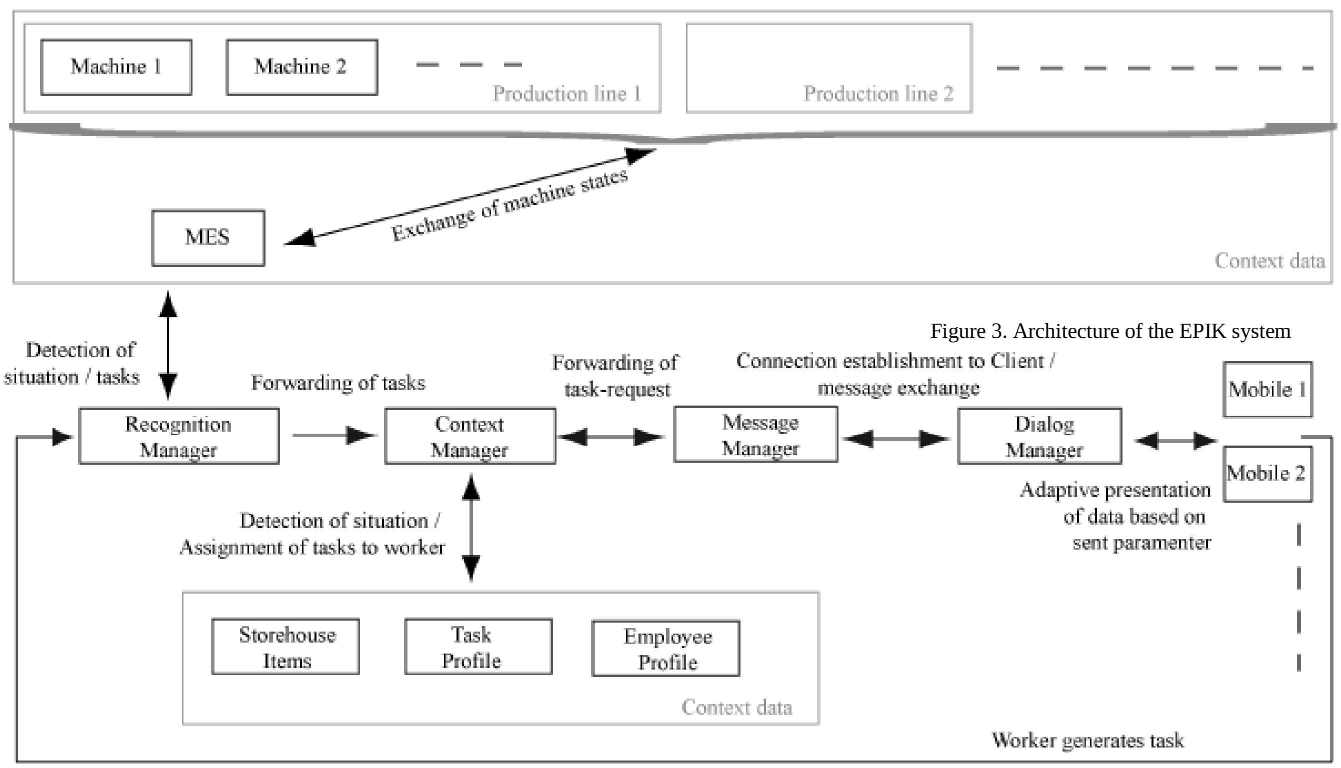


Figure 3. Architecture of the EPIK system

Message Manager: For forwarding the task request and to enable a back channel for the employee a Message Manager is implemented. This manager uses the Extensible Messaging and Presence Protocol (XMPP) for sending and receiving Message. We defined constants that are attached in the subject of the message in order to interpret the intent of the message and thereupon execute the certain EPIK function.

Dialog Manager: Besides improving the production process with individual support of workers with help of

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

mobile devices a further benefit of the EPIK-System are the adaptive user interfaces. Therefore, a Dialog Manager that is implemented on the mobile devices is responsible for best fitting information representation to the user.

IV. Complexitiy, Urgency and Importance of Tasks

During the preparation for the system logic, it became apparent that the following three attributes are necessary for an efficient assignment of tasks to staff:

The **complexity** of tasks describes the range of difficulty to solve the task. Simple actions with low complexity can be accomplished by unskilled workers. This parameter is important for EPIK to appraise, which worker is able to solve the task and how long will the process time approximately be.

Importance differentiates between important tasks and unimportant ones.

Urgency is characterized by time factor. Tasks which have to be done immediately are very urgent.

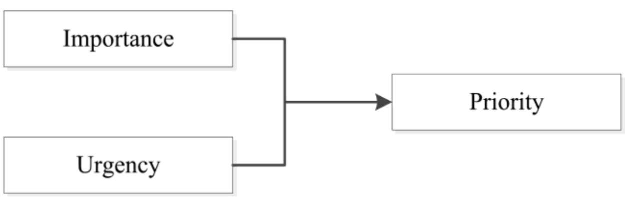


Figure 4. Correlation between importance, urgency and priority

Within this classification, importance and urgency determine when a task has to be done, and in which sequence the several tasks has to be worked on. These two attributes are combined to the **priority**:

- The priority is highest, when the task is urgent and important. These tasks have to be done immediately and other tasks can be interrupted, if they have a low priority.
- Tasks with high importance but low urgency has to be done ongoing and continuously. EPIK is scheduling these tasks ongoing, when there is not a very high priority task.
- High urgency but low importance means that staff with low qualification can do this work immediately.
- Tasks, which are neither important nor urgent have the lowest priority and are scheduled at the end of the tasks.

In this Project, a default set of situations with the associated attributes are specified and declared. For example, an alert is specified as highly important and urgent whereas an upcoming scheduled maintenance is determined as a lower prioritized task. In addition, EPIK provides the functionality for supervisors to parameterize the classification of situations.

The classified tasks are scheduled in the entire task-list of EPIK, which is described more precisely in Section *VII Scheduling*.

V. Assignment of Tasks to Workers

To enhance the efficiency in the production process, a sub-goal of EPIK is to find accurately fitting staff to solve the tasks. To solve this goal, EPIK tries to find the best person for the task and takes different aspects into account:

For the assignment, information of the worker is needed. EPIK collects implicit and explicit data of the worker. For example after the successful completion of an assigned task, EPIK registers the needed time of the worker and increases his qualification value.

Figure 5 displays the selection process for the staff assignment.

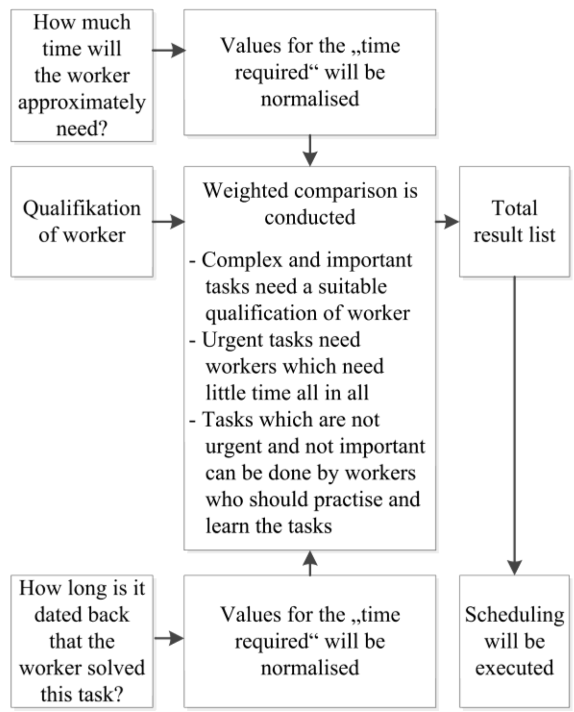


Figure 5. EPIK creates three partial solution lists with workers. Subsequently a weighted comparison is conducted.

The first step is to create three sorted lists which are ordered to the following criteria:

How much time will the worker approximately need? Data of previous tasks from the history are consulted as well as the position of the worker in relation to the location of the machine and task.

The **qualification of the workers** will be compared to the necessary qualification to solve the task. EPIK looks therefore in the profile of the worker and additional in the history of the tasks. The reason for the look in the history is that the workers are increasing their qualification ongoing

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

while using EPIK. Therefore, it is possible that EPIK allocates a task to a worker without the correct qualification, so that the worker enhances his expertise level.

How long is it dated back that the worker solved this type of task? (If the worker never finished this task, the longest duration of the other workers will be used).

The reason to look up the last work on the task is the idea to raise all workers to a similar expertise level. This is to back up redundantly the know how within the company. Therefore, workers with an old date can be preferred within the assignment of tasks to workers. The positive effect is that the workers keep in practice and learned things will not be forgotten immediately.

The final choice of worker is done due to a weighted comparison of the described three lists. The exact weighting can be adjusted individually due to the needs of the company. The weighting factor is differing between the varying tasks.

For very urgent tasks, the factor time is weighted strongly. In contrast therefore, the qualification of the worker is highly weighted, when the task is very important. For little complex and tasks not-important tasks, the thought of training has the dominant weighting. An overview of the weightage is displayed in Figure 6.

Task	Qualification of worker	Approximately required time of worker (normalised)	How long is it dated back that worker solved task (normalised)
High importance	High weighted factor		
High complexity	High weighted factor		
High urgency		High weighted factor	
Low importance			High weighted factor
Low complexity			High weighted factor
	Value worker qualification	Value worker needed time	Value of last task solving
		Result: Qualification + Needet time + Last task solving	

Figure 6. Weighting factors used in EPIK

VI. Scheduling

New tasks will be delegated from EPIK to the best possible worker, even if the worker is still working on another task. Therefore, each worker gets a list of tasks by the time. To ensure that not one worker is always chosen from EPIK (and gets an enormous long list of tasks) and other workers are not contacted at all, EPIK controls by each task assignment the length of the task list of the chosen worker.

Additionally, it is been checked whether the new task has a high priority and where the task should be sorted in within the task-list of the worker. This sorting and allocation of the tasks after priority and length of task-lists is called *scheduling* in EPIK.

The scheduling is working with two different types of task lists:

The **total-task-list** of the whole EPIK system saves all tasks within the production process, which are known already on a stack.

- Tasks, which are already allocated to workers, are marked as allocated and the worker-ID is saved.
- The task list is sorted according to the priority (compare *chapter III complexity, urgency and importance of tasks*). After a defined interval, the total task list is calculated and sorted again. The reason therefore is to avoid that tasks with a low priority won’t be done at all, because they would stay at the last position. With passing time, the tasks are getting higher in the priority.
- Is a task completed, it will be deleted from the total task list and will be saved within the history.

The second list is the **worker-task-list**, which are the individual tasks of the specific worker.

- New tasks are sorted in according to the priority.
- The worker-task-list has a maximum capacity of five tasks.
- Is the length of the list longer than five after sorting in the new task, the last task of the list will be removed from the individual worker-task-list. This task will be marked within the total-task list as not allocated.
  One special case is, when the worker has only high-priority tasks. In this case, the last task will be removed from his list and will be allocated directly to another worker.

The priority is very important for the scheduling, since it is the parameter for the order of the tasks. The priority is very important as well to identify high-priority tasks, which has to be done immediately. To identify these, a threshold is used within EPIK. If the priority of a task is above the threshold, EPIK grabs the attention of the worker to ensure, that he got to know about the task. Furthermore, EPIK controls that the task will be conducted promptly.

VII. Conclusion

The Production process is under continuous change [3] and therefore has to be flexible and dynamic. The competition pressure of internationalization demands efficient systems to afford staff costs in western countries any more. EPIK demonstrates solutions for these needs, but can show only the start of the improvement. Questions like "how can the user-interface support efficiently for often changing international workers" are left by and should be studied more thoroughly. Aspects of staff motivation can be addressed as outlook as well. This means, how can the user interface assist the worker to keep motivated on the job? This question refers a lot to ongoing research projects regarding the user experience of a user interfaces. Up to now, in the production environment, the focus is still based on the usability of the user interface, with the question, "how can the interface assist the user to work faster". The idea of, that the user interface could have such a good user experience, that it makes the user more motivated or happier is in the context of production seldom referred. Palviainen & Väänänen-Vainio-Mattila [6] is one of these references and describes that there is great potential regarding user experience (UX) in the automation industry. Palviainen & Väänänen-Vainio-Mattila point out that "… there are several reasons to assume that the UX paradigm will influence the machinery industry. The level of automation is increasing and the nature of work is shifting from monotonous, low level process control towards expert and team work." Therefore it would be very interesting to make thorough studies to find out how can the user experience affect the efficiency of production.

VIII. Acknowledgments

The project EPIK is financed by the Baden-Württemberg Stiftung gGmbH in the frame of the innovation and research program "Effiziente Produktion durch IKT".

References

[1] Thiel, S. Pfannstiel, S. and Bierkandt, J (2011), "Adaptive Cooperation and Information Management for Efficient Deployment in Production*",* 21st international Conference on Production Research, Stuttgart, Germany, 5pp.,references without page number.

[2] Schlegel, T., and Thiel, S. (2010). „Interaction generation for decentralized process execution*"*. (D. Pham, E. Eldukhri, & A. Soroka, Hrsg.) Dunbeath, Scotland, UK: Whittles Publishing.

[3] Scientific development programm "Research for the Production of Tommorow" of the Federal Government Department of Education and Research in Germany 2012: http://www.bmbf.de/foerderungen/17740.php, retrieved: May, 2012.

[4] Kluge, A. (2004), „Wissenserwerb für das Steuern komplexer Systeme",. Germany: Pabst Science Publishers.

[5] Schöning, U. (2001), „Algorithmik", Berlin: Spektrum, Akad. Verl.

[6] Palviainen, J and Väänänen-Vainio-Mattila, K. (2009), "User Experience in Machinery Automation: From Concepts and Context to Design Implications"; Human Centered Design, Lecture Notes in Computer Science, 2009, Volume 5619/2009, 1042-1051, DOI: 10.1007/978-3-642-02806-9_119

Copyright (c) IARIA, 2012.
ISBN: 978-1-61208-225-7

Window Panes Become Smart

How responsive materials and intelligent control will revolutionize the architecture of buildings

Sotirios D Kotsopoulos, Federico Casalegno
Comparative Media Studies
Mobile Experience Laboratory
Massachusetts Institute of Technology, USA
e-mail: skots@mit.edu, casalegno@mit.edu

Massahiro Ono, Wesley Graybill
Department of Computer Science
Computer Science & Artificial Intelligence Laboratory
Massachusetts Institute of Technology, USA
e-mail: hiro_ono@mit.edu, wgray496@mit.edu

***Abstract*— The use of electronically activated, variable transmission materials, and artificial intelligence control methods in buildings can enhance environmentally and socially sustainable behaviors. It is presented a smart façade, allowing automated control of temperature, daylight, view and privacy in the interior of a prototype house. The programmable materials of the façade allow the precise adjustment of solar transmittance and visibility, while an intelligent controller optimizes the house performance based on the conditions. Using constraint violations as a measure of success, the intelligent controller outperforms the deterministic control models.**

Keywords— ***Responsive materials; intelligent control; smart windows.***

I. Introduction

This paper acknowledges the potential of programmable materials and artificial intelligence (AI) methods to building control, in enhancing the quality of environmentally and socially sustainable living environments. Responsive, interactive skins can bridge the gap between the traditional modes of building and the recent developments in material science, artificial intelligence, and human-computer interaction. The merits from the adoption of technological innovation in buildings can be multifaceted. The ability to monitor and modify the attributes of architectural components can have far reaching implications in energy efficiency and in the ways we practice design. This paper addresses the designing of 'smart windows', where the traditional form and functionality of windows are revisited with a view to integrate the current advances in electrically activated materials research and control systems engineering.

A window is a transparent or translucent opening in a wall, or a door that allows the passage of light and, if not closed, air and sound. Windows are usually glazed or covered in some transparent or translucent material like float glass. They are held in place by frames, which prevent them from collapsing and they can be un-operable or operable. Operable windows may be actuated by the users, to allow ventilation, or they may remain closed, to exclude inclement weather. Windows play important role in the adjustment of the interior lighting and thermal conditions in buildings, and in the relationship between interior and exterior. They also obtain significant social, aesthetic and cultural associations. Primitive windows were simple holes on the wall. Gradually, windows were covered with animal hide, cloth, or wood, and shutters that could be opened or closed were soon added. Over time, materials such as small pieces of mullioned glass, paper, flattened pieces of translucent animal horn, or plates of thinly sliced marble have been used in window panes. Ordinary glass windows became common in homes in the early 17th century, while modern-style floor-to-ceiling windows became possible only after the industrial glass making process was perfected.

Today, responsive materials and intelligent control methods, promise to add new dynamics to windows, including the real time, adaptation of their aesthetic presence and their performance, based on given conditions. Smart windows aim to take active role in the dynamic optimization of building performance with low operational cost. For example, by actively controlling the state of electrochromic glass panes, the solar transmittance of windows becomes a programmable feature, by which it is possible to regulate the interior illuminance and temperature [1], [2]. The transitions of smart windows, optimally operated, can contribute to the reduction of energy consumption by heating and cooling [3]. Further, polymer dispersed liquid crystal films (PDLC) and suspended particle displays, can eliminate the need for traditional blinds and shutter systems and revolutionize building aesthetics. Optical dimming and density variation can replace mechanical actuation with solid-state shading, glare and view-control. Windows equipped with such capacities, can provide new ways to think about the management of energy, of daylight and of privacy, and can drastically transform the way we perceive and inhabit the built environment [4].

An example of the advantages in optimizing these technologies at residential scale is presented next. It is examined a building element of a prototype house –a façade involving a matrix of smart windows – and the adjustment of the lighting and thermal conditions at the house interior. The façade operates as a dynamic filter between exterior and interior. It filters solar radiation and heat by allowing the modification of the chromatism and light transmittance of each individual window. Varying the number and the distribution of the active windows on the façade permits the regulation of the incoming sunlight and heat, and affects the aesthetic presence and performance of the house. The apparatus of the example, yields optimum façade configurations through the efficient management of the electrochemical properties of the windows by an intelligent control system.

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-225-7

Figure 1. South elevation of the prototype house in Rovereto, Italy.

II. BACKGROUND

The connected sustainable home concept – a prototype of which is under construction in Rovereto, Italy (Figure 1) – is an exemplary building structure resulting from the integration of innovative physical and computing architecture. This prototype house constitutes a responsive environment, aiming at remaining well tuned to the comfort levels of the inhabitants, improving quality of life, and supporting socially sustainable living.

One of the most notable features of the house is its south façade. The windows of the façade incorporate electrochromic technology, which permits the regulation of the illuminance and temperature at the house interior through the modification of the chromatism of each individual windowpane. The smart windows are optimally managed by an intelligent control system, aiming to reduce the use of Heating, Ventilation and Air Conditioning system (HVAC) and artificial lighting, and to keep the overall energy consumption of the house minimal overtime. The intelligent control system of the house exploits the passive thermal capacity of the building envelope. During the hot summer days, keeping the interior temperature lower than the exterior becomes a high priority. To protect the interior from direct sun exposure, the control system sets the electrochromic material of the windows to its minimum solar transmittance. Conversely, during the cold winter days, taking advantage of the sun heat becomes a high priority. To expose the interior to the winter sun, the control system sets the electrochromic material to its maximum solar transmittance, thus permitting the storage of natural heat in the home's building structure. Hence, sunlight, heat and shade are used to maintain comfortable interior conditions with minimum use of artificial heating, cooling and lighting.

The key to this approach is the fine management of the passive and active systems of the house. Natural conditions vary and so do the activities of the inhabitants. Programmable materials enabled by the intelligent control, contribute in securing proper living conditions, while keeping energy cost minimal. The residents are prompted to specify the preferred indoor temperature, humidity and illumination, or they can allow the system to take charge, yielding suitable conditions based on the weather. The intelligent control maintains the stability of the indoor environment, whenever the building is occupied. Of course, uncertainty in weather and occupancy patterns poses a risk of failure. The control system limits this risk by explicitly addressing uncertain factors.

After an overview of related background work, this presentation is divided into two main parts, namely, *responsive materials* and *intelligent control*. The first is exposing the performance and the technical characteristics of the responsive materials used in this project. The second is presenting the intelligent control system. More specifically, the setting of a performed simulation test is outlined and the comparison of the proposed controller with two other baseline models is discussed. This presentation concludes with a discussion summarizing the objectives and the contributions of this research.

III. RELATED WORK

Given that electrochromic technology is not broadly used, a number of papers describing the state of the art in this particular research domain are referenced. For example, [1] describes a study in which the effects of electrochromic technology are monitored in a cube 3.0 m x 3.0 m x 3.0 m; [2] presents a technical comparison of data determining the physical features of electrochromic glass; [3] offers an overview on automated lighting and energy control systems; finally, [4] reviews the historical basis of building compliance methods determining minimum energy efficiency and comfort standards, and proposes a new basis for more efficient metrics.

The application of AI methods to building control has been pursued by computational sustainability research. For example, [5] employs the stochastic model-predictive control (SMPC) approach to significantly reduce the energy consumption of a building with stochastic occupancy model. Further, [6] models end user energy consumption in residential and commercial buildings. Another relevant work is [7], which uses a robust plan executive for autonomous underwater vehicles. The proposed plan executive, p-Sulu, is built upon the Iterative Risk Allocation algorithm [8] and a deterministic plan executive, Sulu [9].

IV. RESPONSIVE MATERIALS

The smart façade of the prototype house is a matrix of programmable windows that are individually addressable. It was designed to achieve three performative objectives: (1) regulate the percentage of sun-heat that penetrates the house; (2) regulate the interior illuminance; and (3) regulate the airflow. Each window pane is an overlay of two electronically switchable materials. The first layer, the electrochromic glass, is applied on the external glazing to provide the desirable degree of sunlight penetration securing daylight and thermal performance. The second layer, the polymer dispersed liquid crystal film (PDLC), is applied on the internal glazing to provide the desirable degree of visibility, securing privacy (Figure 2).

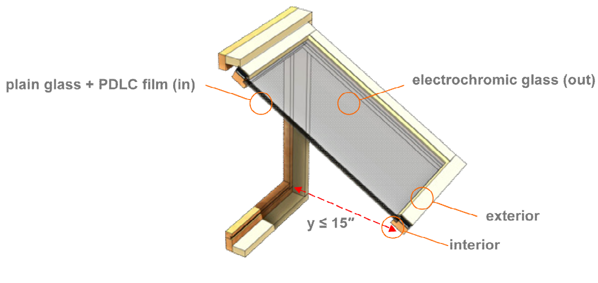


Figure 2. Axonometric section of the window. Each window pane is an overlay of two materials: electrochromic glass and PDLC film.

Copyright (c) IARIA, 2012.

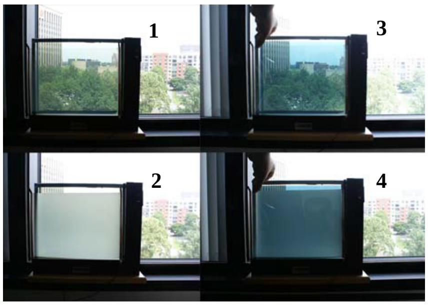


Figure 3. The overlay of materials: (1) the electrochromic and the PDLC layers are inactive; (2) the electrochromic layer is innactive and the PDLC layer is active; (3) the electrochromic layer is active and the PDLC layer is innactive; (4) the electrochromic and the PDLC layers are active.

Smart windows operate in selective mode, admitting natural light, heat, and air, or excluding any of the above as needed. At the global level, the windows are directed by the central control system, without having any interference among them. At the local level, each window is driven by its own low level intelligence, software and custom electronics that enable the activation of its switchable materials. Each window pane is equipped with a photocell to calculate the amount of light that it is exposed to, while an IR sensor detects the presence of the residents. The windows provide this sensory feedback to the central control that compiles the information to activate any number of windows is required. Since the switchable materials have varying response times and exhibit different optical, thermal and power consumption characteristics, their activation processing is preplanned. The slow dimming response of 8 minutes, of the electrochromic glass, is suitable for controlling sunlight and heat, while the instant transition of the PDLC film is useful for controlling shade and privacy. The PDLC film can be switched to entirely block the view, or it can adjust the degree of its opacity based on the detection of specific patterns of user activity. For example, the PDLC film can be programmed to output a varying opacity, corresponding to the gesture of holding the hand in front of the photocell for a range of time durations.

The management of the façade is driven by the autonomous control system of the house. The electrochromic technology operates as an alternative to a traditional screening system [10]. Façade patterns can be enabled to adjust the transmittance value of window panes. The values of transmittance (τ) vary from 60-75%, for idle glass, to 3-8% for active obscured glass [2]. Extensive analysis and evaluation of simulation data allowed to discover how the activation of the electrochromic glass affects the levels of thermal comfort and interior daylight illuminance, on the specific building, at the specified site, throughout the year. The simulation software used in this analysis was Relux Professional by Relux Informatik AG. The climatic data used in the simulations was provided by the database of the software, which also integrates the lighting standards determined by Italian law. Relux Vision (a plug-in of Relux Professional) was used in the ray-tracing simulation. Typical outputs of each simulation test included: a) the min/max and average values of illuminance and b) the lighting uniformity values and average daylight factor.

Illuminance in daylight conditions is subject to constant change of intensity and distribution. Interior daylight is affected by the movement of the sun and of the clouds in the sky. The parameters determining the optimum daylight vary based on the daytime and month of the year. The existing simulation models for buildings provide a poor account of daylight modification [3]. These models identify and approximate uncomplicated settings, which do not reproduce the natural sunlight conditions. For our purpose, two models – established by the *Commission Internationale de l'Eclairage* (CIE) – were used in the analysis and evaluation of natural light. The first, is the Standard Overcast Sky model, which provides an account of the sunlight emitted through cloudy sky. The second is the Standard Clear Sky model, which computes sunlight under the assumption that the sun is the single lighting source, without calculating the diffused and reflected light by the sky. Although the two models offer an accurate account of the interior sunlight distribution in two distinct settings, they provide only a partial understanding of the phenomenon of daylight change. Because, they exclude the assessment of transitory conditions, such as the momentary passage of clouds across the sun [11].

Despite these limitations, the aim of calculating the optimum façade settings is within reach. Exhaustive simulation tests were performed and data corresponding to the daylight conditions for each day of the year, was collected, classified and used for the integration of an optimization algorithm into the control system. After setting a desired value of interior daylight illuminance in lux the simulation tests helped to identify the required numbers of active windows to reach this value for every day of the year. A predictive model associating coverage ratio and illuminance, in any specified time interval was extracted and integrated into the control system. Based on this model the number of active windows ensuring the illuminance threshold in overcast sky throughout the year, ranges from 50 to 75, in a total of 100.

The produced database was used as a foundation for optimizing sunlight and heat. The combination of this information with real time feedback from sensors enables the reprogramming of the façade, even in conditions that cannot be predicted by the static simulation model. The overall system achieves two complementary objectives: a) determines the optimum number of active electrochromic glass units to reach the desired interior temperature and illuminance level, and b) confines the current façade patterns to this threshold.

	1p.m. -21st	x	α
1	December	0	0,00
2	January	18	0,18
3	February	45	0,45
4	March	62	0,62
5	April	72	0,72
6	May	75	0,75
7	June	75	0,75
8	July	75	0,75
9	August	72	0,72
10	September	63	0,63
11	October	45	0,45
12	November	16	0,16
13	December	0	0,00

Table 1. Maximum numbers of active windows throughout the year at 1 p.m. in *Standard Overcast Sky* after simulation with Relux.

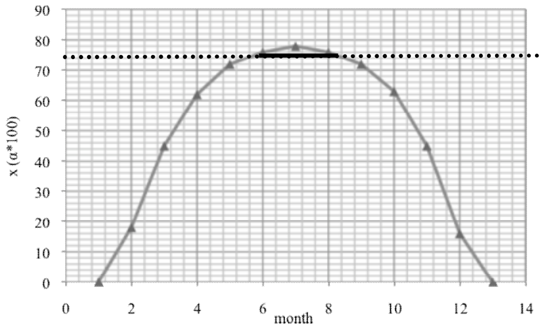


Figure 4. A model gives the maximum number of active windows yielding the threshold value during the whole year in Overcast Sky. In doted line is the 75 % limit.

The interior daylight simulations confirmed that the façade could switch from one pattern to another while maintaining the same activation ratio, in order to possibly satisfy other factors of performance. A generative grammar producing any class of patterns was then specified so that any desired level of interior daylight conditions could be reached. In [12] the simulation process of the electrochromic windows is described in detail and the generative grammar capturing the language of the optimum façade configurations for the prototype is presented.

V. INTELLIGENT CONTROL

The next sections describe the computational apparatus managing the states of the responsive materials and the temperature at the house interior. The windows are controlled in a concerted manner by the central control without having autonomous intelligence capability. The controller is implemented in C++. Uncertainty in outdoor temperature is taken into account. The hardware hosting the controller is a standard computer with 8GB of RAM and Intel Core i7 processor. A Mini-ITX system secures low energy profile. The controller also manages illumination and humidity. These extensions, which are included in the prototype, are not presented here. First, the three main control requirements enabling this overall achievement are outlined, and then, the proposed solution is exposed.

A. Goal-directed planning with continuous effect

The controller should allow residents to specify desired ranges of room temperature (i.e., state constraints in a continuous domain) as well as their schedule (i.e., temporal constraints). Concisely, the control must be able to execute plans with time-evolved goals, which are specified as a sequence of state and temporal constraints. Then, it must optimally adjust the operation of windows and the HVAC system, so that the specified constraints are satisfied.

B. Optimal Planning

While guaranteeing that the time evolved goals are achieved, the controller should also minimize the use of non-renewable energy consumption.

C. Robust planning with risk bounds

Optimal plan execution is susceptible to risk when uncertainty is introduced. The house, involves a risk of failure to maintain the room temperature within a specified range due to unexpected climate changes. In the winter, when the residents are absent, the energy consumption can be minimized by turning off the heating. But, this involves a risk that the pipes may freeze. Such risks must be limited to acceptable levels specified by the residents. The plan executive guarantees that the system is able to operate within these bounds. Such constraints are called chance constraints.

D. Proposed Solution: p-Sulu

The proposed solution [13] to overcome these challenges is a newly developed robust plan executive called p-Sulu. Previous research [14] presents the risk-sensitive finite-horizon planner, called p-Sulu planner, which has the three capabilities. However, p-Sulu planner is an off-line planner, meaning that it pre-plans the control sequence for the entire plan duration. The off-line planner could not possibly control the Sustainable Connected Home due to two technical challenges. First, the house must be operated continuously with no interruptions, while p-Sulu planner can plan only for a finite time duration. Second, building control requires frequent re-planning every few seconds, while p-Sulu planner's solution time is typically on the order of minutes. For example, a simple path planning problem with only 10 time steps takes about 30 seconds to solve [14]. Moreover, the computation time of p-Sulu planner varies widely from problem to problem.

To overcome the first challenge a receding horizon control approach was adopted. At each planning cycle, a planning problem is solved with a finite duration, which is called a horizon. In the next planning cycle, the planning problem is solved again over a horizon with the same duration starting from the current time (hence, the horizon is “receding”), by considering the latest observation of uncertain parameters. This re-planning process is repeated with a fixed time interval. The second challenge was faced by building upon an anytime algorithm for chance-constrained programming, called Iterative Risk Allocation (IRA) [8]. IRA was originally developed for path planning problems without time-evolved goals [8]. The proposed approach extends IRA to deal with time-evolved goals.

p-Sulu RH takes as an input a plan representation called a *chance-constrained qualitative state plan (CCQSP)*, which encodes both time-evolved goals and chance constraints [14]. The goal is to find a control sequence, which is an assignment to real-valued control variables, as well as a schedule, which is an assignment of discrete execution time to events.

An advantage of p-Sulu over existing deterministic plan executives is that it can explicitly consider a stochastic plan model, which specifies probabilistic state transitions in a continuous domain in the following form:

$$x_{t+1} = A_t xt + B_t u_t + w_t$$

where x_t is a continuous state vector at time t, u_t is a continuous control vector at t, and w_t is a disturbance whose probability distribution is known.

A chance-*constrained qualitative state plan* (*CCQSP*) is a four-tuple $P = \langle \mathcal{E}, \mathcal{A}, \mathcal{T}, \mathcal{C} \rangle$ where $\mathcal{E}$ is a set of discrete events, $\mathcal{A}$ is a set of *episodes*, $\mathcal{T}$ is a set of *temporal constraints*, and $\mathcal{C}$ is a set of *chance constraints*.

An example of a resident schedule for a day is presented in Figure 5 as a *chance constrained qualitative state plan* or *CCQSP*, while the state space of the same graph appears in Figure 6.

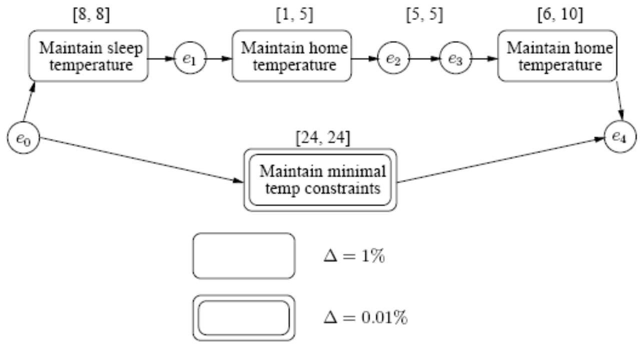


Figure 5. An acylic directed graph depicting the resident's schedule in the sample planning problem for the Connected Sustaianble Home

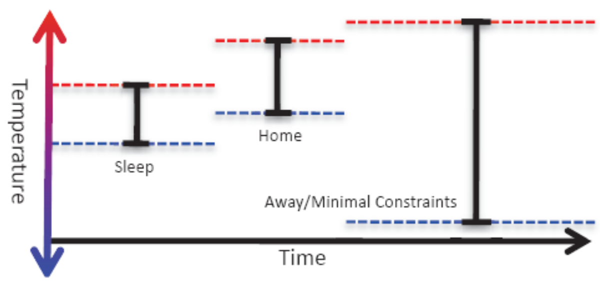


Figure 6. The state space of the previous graph based on the 3 specified ranges Home, Asleep, and Away.

The schedule is described first in plain English, as follows:

"Maintain a comfortable sleeping temperature until I wake up. Then, maintain room temperature until I go to work. I may work at home, but I have to do 5 hours of work at the office sometime between 9 am and 6 pm. No temperature constraints while I am away. When I get home, maintain room temperature until I go to sleep. The probability of failure of these episodes must be less than 1%. The entire time, make sure the house doesn't get so cold that the pipes freeze. Limit the probability of such a failure to 0.01%."

At any time, the schedule confines a temperature range to maintain over some duration. In the experiments, it was assumed that a resident can specify one of 3 ranges: Home, Asleep, and Away. In actuality, one is able to select any number of temperature ranges. It was assumed that the temperature must remain between 20° and 25° C while the resident was at Home, between 18° and 22° C while Asleep, and between 4° and 35° C while Away, to ensure that the pipes would not freeze. Home and Asleep episodes, were associated to a single chance constraint class, with risk bound 10 %. This is the risk the resident is willing to take that the temperature may become uncomfortable. Away episodes were associated to a single chance constraint class with risk bound 0.01 %. This is the acceptable risk by the resident that the pipes may freeze.

The above schedule was tested on a simulation setting [15]. Two baseline models were compared with the proposed controller p-Sulu: (a) a PID was used to illustrate the energy savings of a model-predictive controller compared to a traditional heating controller, and (b) Sulu, the deterministic predecessor to p-Sulu, was used to illustrate the robustness of p-Sulu. The initiating stage of the PID controller was 21° C. The planning was based on the resident schedule graph spanning a week. Then, uncertainty w_t is introduced and the plan is executed in a stochastic simulation. Each controller was evaluated on the basis of energy savings and percentage of executions that fail because of constraint violations.

The next figures illustrate the results of a stochastic simulation over two different days in the year, January 1 (Figure 7) and July 1 (Figure 8). It should be noticed that the deterministic predecessor Sulu of the newly proposed controller plans right up to the edge of the constraints, often violating constraints when uncertainty is introduced, while p-Sulu leaves a margin of the stochastic simulation.

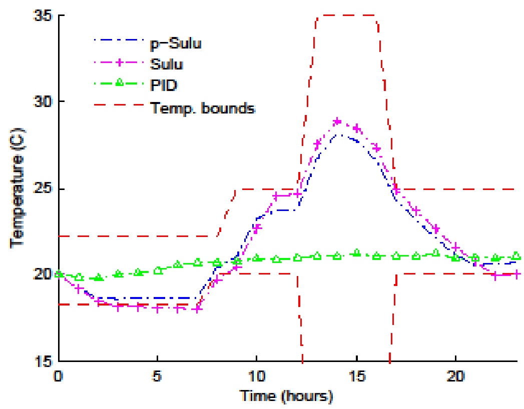


Figure 7. Results for PID, Sulu, and p-Sulu controllers on January 1. p-Sulu controller appears in blue line.

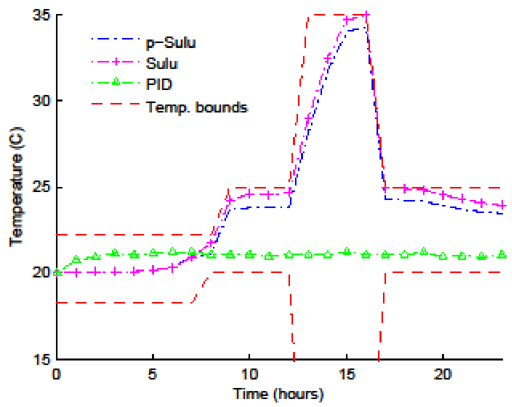


Figure 8. Results for PID, Sulu, and p-Sulu controllers on July 1. p-Sulu controller appears in blue line.

	Winter		Summer	
	Energy	Violation Rate	Energy	Violation Rate
p-Sulu	1.9379×10^4	0.000	3.4729×10^4	0
Sulu	1.6506×10^4	0.297	–	–
PID	3.9783×10^4	0	4.1731×10^4	0
	Spring		Autumn	
	Energy	Violation Rate	Energy	Violation Rate
p-Sulu	3.3707×10^4	0	3.8181×10^4	0
Sulu	3.0954×10^4	0.308	3.6780×10^4	0.334
PID	3.9816×10^4	0	3.9955×10^4	0

Table 2. Comparison of energy use and failure rate of the PID, Sulu, and p-Sulu controllers over a week-long schedule in all four seasons. Failure rate is measured as the percentage of time steps with constraint violations.

Table 2 presents the results on the weeklong scenario, averaged over 100 Monte Carlo trials each. Notice that the failure rate of p-Sulu is actually significantly lower than the risk bound. Out of the 200 trials of Sulu in the winter 159 failed to complete due to infeasibility. Of the 41 trials that completed, every single one had constraint violations on at least 31 of the 168 time steps of the plan. All trials of Sulu in the summer failed due to infeasibility.

Furthermore, the results illustrate that although the PID controller did not violate temperature constraints, the Sulu and p-Sulu performed drastically better in energy consumption. While p-Sulu sacrificed some energy savings compared to Sulu, these loses were met with drastic improvements in robustness.

As a measure of comfort, one may compare the trials that completed and consider the fraction of time steps on which a constraint was violated. Table 2 shows that out of the 168 time steps, Sulu violated constraints on 29.7% of time steps in the winter, 30.8% in the spring, and 33.4% in the autumn. Only a single trial of p-Sulu violated a single constraint in winter. All other seasons, all trials satisfied all temperature constraints. Using constraint violations as a measure of success, the approach of p-Sulu outperforms the deterministic approach of Sulu on every trial. Averaged across all trials, p-Sulu exhibits a difference of 30.88% in improvement in comfort over Sulu.

These results illustrate how critical risk-sensitive control is in guaranteeing resident comfort and encouraging the adoption of the technology. A control system that produces uncomfortable conditions 30% of the time would quickly be abandoned by the users.

VI. Discussion

Programmable materials and AI methods to building control can be used to enhance sustainable living in buildings. This paper addresses the designing of smart windows at residential scale. The aim and functionality of windows are revisited with a view to integrate innovative building materials and control solutions.

The features of a programmable façade for a prototype house were discussed. These features, rest on the capacity to monitor and modify the state of each individual window in real time. Each windowpane is an overlay of two electronically switchable, variable transmission materials. The first, electrochromic layer, provides the desirable degree of sunlight penetration, securing sunlight and thermal performance. The second, polymer dispersed liquid crystal (PDLC) layer, provides the desirable degree of visibility, securing privacy. At the global level, the windows are directed by a central control system, without having interference among them. At the local level, each window is using low-level intelligence, software and custom electronics that enable the activation of its switchable materials.

An autonomous control system aiming at optimizing the long-term energy performance of the house enables the management of the programmable materials of the façade. The control compiles data related to the seasonal levels of light and heat, and real time feedback from sensors, to reprogram the state of the façade, in order to take advantage of the sun and to minimize energy use. Hence, the façade functions as a responsive component, admitting natural light, heat, and air, as needed. And further, the adjustments of the façade transform how the house is perceived from the public street.

Windows are typically operable building components that are crucial to the good thermal performance of buildings. Windows had evolved from holes covered with cloth, mullioned glass, or paper, to modern-style floor-to-ceiling curtain walls, sealed with industrial glass. Today, responsive materials and AI control methods promise to add new dynamics to the functionality and the aesthetics of windows.

The "smart window" offers a new basis to rethink the role of windows in buildings. It combines environmentally responsible design with inventive technology, and it points to new ways of improving the interaction among people. The reprogramming of a smart window enables autonomous, responsive and interactive behaviors, while taking into account not only functionality, but also social and cultural aspects. Smart windows re-establish some familiar attributes of traditional windows in the digital age.

For example, modern-style floor-to-ceiling glass curtain walls backed with constant artificial lighting, heating, ventilating and air-conditioning, were aiming to secure air-and-sound sealed interior environments. These conventions of performance no longer apply, since the use of artificial cooling and heating is energy intensive. At the social level, modern curtain walls were restraining human behavior by confining people strictly "inside" or "outside" of a building.

The flexible modes of use of the dynamic façade range from moderation of view and air for each window, to privacy control, personalized daylight or shade, and ultimately personal communication and self-expression. Except from its function as interior climate regulator, the façade operates as a responsive mediator between private and public domain.

It is likely that in the future, interactive building components, will give rise to new conceptions of space, and eventually to "new architectures", where people will become engaged with their environments in novel emotional, social and intellectual modes.

Acknowledgment

This research was conducted within the *Green Connected Home Alliance* between the Mobile Experience Lab, at the Massachusetts Institute of Technology and the Fondazione Bruno Kessler in Trento, Italy

Figure 9. Visualization of electrochromic patterns on the south façade of the prototype house, as percieved from the exterior and the interior.

REFERENCES

[1] E. S. Lee, D. L. Di Bartolomeo, J. H. Klems, M. Yazdanian and S. E.Selkowitz, 2006, "Monitored energy performance or electrochromic windows for daylighting and visual comfort", ASHRAE Summer Meeting, Quebec City, Canada.

[2] T. Hausler, U. Fischer, M. Rottmann and K. H. Heckner, 2003, "Solar optical properties and daylight potential of electrochromic windows, International Lighting and Colour Conference, Capetown.

[3] S. Selkowitz, O. Aschehoug and E. S. Lee, 2003, "Advanced interactive facades – critical elements for future buildings?", presented at USGBC expo.

[4] J. Mardaljevic, L. Heschong and E. S. Lee, "Daylight metrics and Energy savings", Lighting Research and Technology, 2009 41:261.

[5] A. E. D. Mady, G. M. Provan, C. Ryan, and K. N. Brown, 2011. "Stochastic model predictive controller for the integration of building use", Proceedings of Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI-11), Special Track on Computational Sustainability and AI.

[6] J. Z. Kolter, and J. Ferreira, 2011, "A large-scale study on predicting and contextualizing building energy usage", Proceedings of Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI-12), Special Track on Computational Sustainability and AI.

[7] H. Li, and B. C. Williams, 2011, "Hybrid planning with temporally extended goals for sustainable ocean observing", Proceedings of Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI-11), Special Track on Computational Sustainability and AI.

[8] M. Ono, and B. C. Williams, 2008, "An efficient motion planning algorithm for stochastic dynamic systems with constraints on probability of failure", Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (AAAI-08).

[9] T. Leaute, and B. C. Williams, 2005, "Coordinating agile systems through the model-based execution of temporal plans", Proceedings of the Twentieth National Conference on Artificial Intelligence (AAAI).

[10] F. Gugliermetti, and F. Bisegna, 2003, "Visual energy management of electrochromic windows in Mediterranean Climates", Building Environment, 479-492.

[11] J. Wienold, 2007, "Dynamic simulation of blind control strategies for visual comfort an energy balance analysis", International Building Performance Simulation Association, Beijing, China, 1197-1204.

[12] S. Kotsopoulos, G. Cara, W. Graybill, and F. Casalegno, 2012, "The dynamic facade pattern grammar", Massachusetts Institute of Technology.

[13] M. Ono, W. Graybill, and B. C. Williams, 2012, Risk-sensitive plan execution for Sustainable Connected Home, Proceedings of AAAI Conference on Artificial Intelligence (AAAI-12), Special Track on Computational Sustainability and AI, (*in* review)

[14] M. Ono, 2012, Robust, Goal-directed Plan Execution with Bounded Risk. Ph.D. Dissertation, Massachusetts Institute of Technology.

[15] W. Graybill, 2012, Robust, Goal-Directed Planning and Plan Recognition for the Sustainable Control of Homes. Master's Thesis, Massachusetts Institute of Technology.