# AMBIENT 2012

The Second International Conference on Ambient Computing, Applications, Services and Technologies

ISBN: 978-1-61208-235-6

September 23-28, 2012

Barcelona, Spain

**AMBIENT 2012 Editors**

Maarten Weyn, Artesis University College of Antwerp, Belgium

Pascal Lorenz, University of Haute Alsace, France

# AMBIENT 2012

# Forward

The Second International Conference on Ambient Computing, Applications, Services and Technologies (AMBIENT 2012), held on September 23-28, 2012 in Barcelona, Spain, was devoted for a global view on ambient computing, services, applications, technologies and their integration.

On the way for a full digital society, ambient, sentient and ubiquitous paradigms lead the torch. There is a need for behavioral changes for users to understand, accept, handle, and feel helped within the surrounding digital environments. Ambient comes as a digital storm bringing new facets of computing, services and applications. Smart phones and sentient offices, wearable devices, domotics, and ambient interfaces are only a few of such personalized aspects. The advent of social and mobile networks along with context-driven tracking and localization paved the way for ambient assisted living, intelligent homes, social games, and telemedicine.

The conference provided a forum where researchers were able to present recent research results and new research problems and directions related to them. We welcomed technical papers presenting research and practical results, position papers addressing the pros and cons of specific proposals, such as those being discussed in the standard forums or in industry consortiums, survey papers addressing the key problems and solutions on any of the above topics, short papers on work in progress, and panel proposals.

We take here the opportunity to warmly thank all the members of the AMBIENT 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 that dedicated much of their time and efforts to contribute to the AMBIENT 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 also gratefully thank the members of the AMBIENT 2012 organizing committee for their help in handling the logistics and for their work that is making this professional meeting a success. We gratefully appreciate to the technical program committee co-chairs that contributed to identify the appropriate groups to submit contributions.

We hope the AMBIENT 2012 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in ambient computing research.

We hope Barcelona provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

**AMBIENT 2012 Chairs**

**AMBIENT General Chair**
Maarten Weyn, Artesis University College of Antwerp, Belgium

**AMBIENT Advisory Chairs**
Yuh-Jong Hu, National Chengchi University-Taipei, Taiwan (R. O. C.)
Naoki Fukuta, Shizuoka University, Japan
Andrzej Skowron, Warsaw University, Poland

**AMBIENT Research Liaison Chairs**
Rémi Emonet, Idiap Research Institute, Switzerland
Mieczysław A. Kłopotek, Polish Academy of Sciences - Warszawa, Poland

**AMBIENT Special Area Chairs**

**Agents**
Joseph Giampapa, Carnegie Mellon University, USA
Adina Magda Florea, University "Politehnica" of Bucharest, Romania

**Context-awareness**
Sylvain Giroux, Université de Sherbrooke, Canada

# AMBIENT 2012

## Committee

**AMBIENT General Chair**

Maarten Weyn, Artesis University College of Antwerp, Belgium

**AMBIENT Advisory Chairs**

Yuh-Jong Hu, National Chengchi University-Taipei, Taiwan (R. O. C.)
Naoki Fukuta, Shizuoka University, Japan
Andrzej Skowron, Warsaw University, Poland

**AMBIENT 2012 Research Liaison Chairs**

Rémi Emonet, Idiap Research Institute, Switzerland
Mieczysław A. Kłopotek, Polish Academy of Sciences - Warszawa, Poland

**AMBIENT 2012 Special Area Chairs**

**Agents**
Joseph Giampapa, Carnegie Mellon University, USA
Adina Magda Florea, University "Politehnica" of Bucharest, Romania

**Context-awareness**
Sylvain Giroux, Université de Sherbrooke, Canada

**AMBIENT 2012 Technical Program Committee**

Sachin Kumar Agrawal, Samsung - Noida (Delhi NCR), India
Eugenio Aguirre Molina, Universidad de Granada, Spain
Youakim Badr, INSA - Lyon, France
Flavien Balbo, University of Paris Dauphine, France
Stephen Barrass, University of Canberra, Australia
Orlando Belo, University of Mihno, Portugal
Daniel Berckmans, Katholieke Universiteit Leuven, Belgium
Stephan Böhm, Hochschule RheinMain, Germany
Philip J Breedon, Nottingham Trent University, UK
Sybille Caffiau, Université Joseph Fourier, Grenoble I, France
Valerie Camps, Universite Paul Sabatier - Toulouse, France
Carlos Carrascosa, Universidad Politécnica de Valencia, Spain
John-Jules Charles Meyer, Utrecht University, The Netherlands
Cezar Collazos, Universidad del Cauca, Columbia
Tim Dams, Artesis University College, Belgium

**Copyright Information**

## Table of Contents

# *TeleMedSys - Web-based Telemedicine System to Support Diagnosis in Rural Areas*

Zoila Gonzalez Videuax
Departamento de Informática
Universidad de Oriente (UO)
Santiago de Cuba, Cuba
zgv@fie.uo.edu.cu

Reinhard Langmann
Competence Center Automation Düsseldorf (CCAD)
Düsseldorf University of Applied Science (DUAS)
Düsseldorf, Germany
langmann@ccad.eu

***Abstract*— This article describes the concept for a web-based telemedical system for use in rural areas with limited communication infrastructures and a lack of medical professionals. The TeleMedSys consists of a patient management system and a medical device system that are available to the doctor and the patient-side assistant during treatment via an Internet browser. The combination of local and remote web applications enables a seamless integration of the medical devices required at the patient's location into the doctor's remote station based on browser technology. The implementation of this system represents an improvement in the living conditions of populations in rural areas.**

***Keywords-Web-based telemedicine; remote medical device; distributed medicine system; web technology; rural telemedicine.***

## I. BACKGROUND AND STATE OF THE ART

Significant efforts are currently underway in various Latin-American countries (e.g., Ecuador, Cuba, Mexico, Brazil) to improve the healthcare situation in hard-to-access, rural and underdeveloped regions; using telemedical solutions.

The Foundation for Telemedicine and eHealth in Ecuador (FUNDETEL) [26] is an example of an institution that together with local universities, is currently working on the collaboration and promotion of telemedicine and telehealth in Ecuador and internationally, not only as a teleconsultation and second opinion tool for rural areas such as the ANDES and COASTAL AREAS, but also in the area of training / online education in telemedicine, participating in the IDBA project for Latin-American Politics and Standards for Telemedicine and Telehealth.

In highly developed industrial countries with an extensive medical infrastructure, telemedical systems are mainly used to improve the quality of life for older people [1][2], as well as in emergency healthcare systems [2][3], and for systems used in particular situations (on ships [5], in the space flight industry [6][7], and in military applications [8][9]).

The available telemedical systems are not suited for regular and conventional treatment processes with only untrained or semi-skilled medical staff available at the patient's location. In addition, the costs of conventional systems available on the market by far exceed the cost limit for practicable systems in developing countries. The technical solutions are based on proprietary and supplier-specific components, while ICT standard technologies are hardly used.

Some web-based telemedical projects planned to be deployed in underdeveloped regions and countries already exist [10][11][12]. However, the respective concepts in most cases only address the required telemedicine, while not considering the actually required telemedical stations, for which conventional systems available on the market shall be used. Further interesting solutions and discussions about telemedicine concerning the topic of this paper you can find in [14-25].

In rural and isolated areas with geographical limitations, telemedical stations that take into account satellite connectivity, power supply such as solar panels, and ready-to-use teleconsultation equipments is urgently needed and not easily available or affordable.

Furthermore, in the year 2008, during the REV conference in Düsseldorf, Maria Teresa Mijares presented a keynote speech [27] on the benefits and advances of telemedicine. This sparked the interest of many of the institutions and universities present at the event, in telemedicine.

Duesseldorf University of Applied Sciences / Competence Center Automation Duesseldorf (CCAD) together with Universidad de Oriente / Centro de Neurociencias Procesamiento de Imagenes y Señales (CENPIS) / Santiago de Cuba decided to research and create a telemedical system (short: TeleMedSys), as an important and affordable tool for the application of telemedicine in rural and isolated areas.

The main problem of this research is a lack of a control system that allows doctors to take care of patients from different rural areas directly from hospitals.

The current status of TeleMedSys is a demonstration example (prototype) that is analyzed and discussed in the present article.

In the next sections of this paper we described different aspects of TeleMedSys. First a deep analysis of the main Requirements and objetives. The Scientific and Technical Solution for the prototype is also made, here is exposed the equipment structure of the TeleMedSys, the functional structure, problems of data analysis and the design and implementation. We finish this paper with the proposal for the use of this system in other environments after finish with first version.

 ISBN: 978-1-61208-235-6 

## II. Requirements and objectives

The main objective of TeleMedSys is to support telemedicine and the remote diagnosis of people living in rural areas, from the doctor stations. TeleMedSys respects the natural and economic conditions of each area and works in dependence of it.

This prototype is not supposed to be a technological innovation; we were trying to solve a critical problem in rural areas using different technologies, particularly Web-technologies.

The main requirements and functionalities of the system are listed below. They were defined in agreement with FUNDETEL for the new telemedical system (TeleMedSys):

- A doctor station (DS), which will be able to cooperate with n patient stations (PS) respectively, will be set up.
- The PS deployment site has no power supply or communication infrastructure. The TeleMedSys will be transported through harsh terrain to the deployment site.
- Only medical assistants will be available at the PS site, which will have to be guided by the doctor during the treatment process.
- The TeleMedSys shall be web-based in order to integrate simple e-learning elements in parallel (work-integrated learning).
- 95% of the treatment process is performed asynchronously between the patient/assistant and the doctor while the remaining 5% are performed synchronously via video conference and/or real-time chat.
- The weekly period of use for a PS amounts to 3-4 hours. The maximum distance between the PS and the respective DS is 2,000 km.
- The medical examination equipment is connected to the PS; it will be unlocked by the DS and operated by the assistants according to the doctor's instructions. The measuring data are automatically provided to the DS.
- The acquisition costs of the TeleMedSys should not exceed EUR 20,000. The operating costs should be less than EUR 1,000 per month.

A particular challenge is the asynchronous treatment process, because in traditional medical practice such a procedure is not required and unknown.

The Düsseldorf University of Applied Sciences initiated a university-internal research project with the involvement of the Universidad de Oriente in Santiago de Cuba for developing the first prototype in 2011.

## III. Scientific & Technical Solution

### A. *Equipment structure of the TeleMedSys*

The device-related part of the DS can be realised with a standard computer, as the physician has the required infrastructure (fixed power supply, Internet access) and no medical devices need to be connected. The development of the PS was the biggest project challenge. Based on the above-mentioned requirements, the PS will therefore be equipped with the following components:

- A robust notebook for use under critical environmental conditions (ruggedized).
- A solar power station featuring a buffered battery operation for the power supply (solar output: 28W; battery: 15Ah, 12,8V; module output: 15V, 1800mA).
- Satellite modem (Inmarsat satellites) for setting up Internet access with 384 Kbit/s download and 240 Kbit/s upload.
- Medical examination devices with Bluetooth interface. The prototype, for the time being, will feature a blood pressure measuring device and a 6-channel ECG device.

Figure 1 shows a picture of the PS device components.

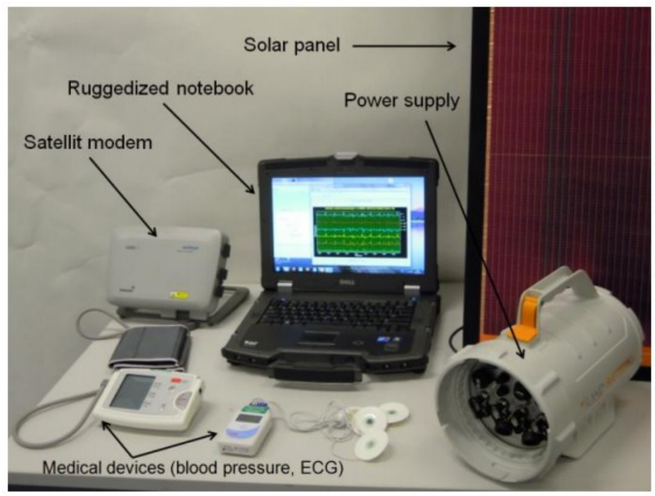



Figure 1. Device components of the Patient Station for the TeleMedSys

The appropriate technology devices will be implemented in accordance with the primary health needs and common diseases of the area. For further development the system should be connected to at least the following additional equipment for medical examinations:

- Ultrasonic diagnostic device;
- Digital camera;
- Body temperature measuring instrument;
- Spirometer.

For all equipment components of the PS, the development of a special shipping box that is suitable for transport over rough terrain is also being planned. After a field test of the prototype, we will also re-examined the appropriateness of the used medical devices with Bluetooth connectivity. Although the wireless connection of the medical device with the PS is a flexible and elegant solution, all units require an extra power supply with batteries, which may in some cases not be available. Moreover, for the communication interface of medical devices (Bluetooth, USB or serial), the respective communication protocols and APIs are required. This is usually not the case, since the devices commonly come as a complete application solution. Exceptions are the devices of the German company Corscience [28] which are used in the prototype of the PS.

The serial protocol is published by Corscience and can be used very well for development purposes. The protocol is designed for simple and memory-saving implementation in a microcontroller. The overhead was kept as low as possible. The principle of this protocol is basically modeled on the PPP (point-to-point protocol), which is often used to establish modem connections. Furthermore, escape sequences are used to filter out reserved bytes (start flag, end flag, escape flag) from the data stream.

### *B. Functional structure*

The TeleMedSys software consists of the following three key components:

- Patient Management System (PMS),
- Medical Device System (MDS), and
- Online Collaboration System (OCS).

The prototype requires the PMS and the MDS (only for asynchronous operation). Later, the prototype will be extended with the OCS.

Figure 2 depicts the functional software structure of the TeleMedSys.

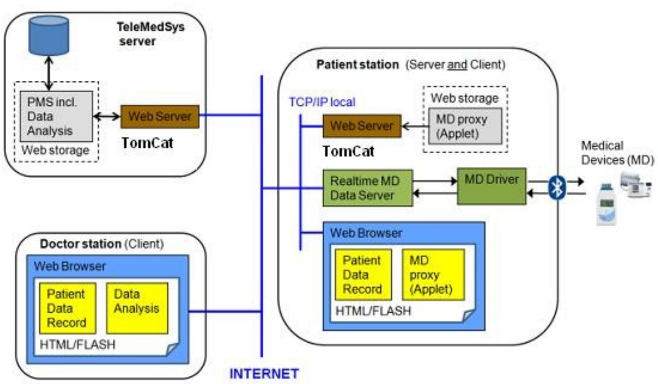



Figure 2. Functional structure of the TeleMedSys

All patient information is stored by the Patient Management System (PMS) on a TeleMedSys server. Both the DS and the PS (doctor and assistant) have access via a web browser to the PMS database on the TeleMedSys server and can edit the information in a Patient Data Record (PDR). The DS is working as a client and uses only a web browser.

The asynchronous patient treatment is the main issue of the PMS (see Figure. 3). It requires separating the overall treatment session for a patient into individual steps for the assistant at the PS and the doctor at the DS. These individual steps are performed alternately by the assistant and the doctor. The system automatically displays the status of all patient sessions and provides information about the next required step for each patient through an event screen.

In principle, the treatment is done in the asynchronous mode of operation similar to the remote diagnostic and service procedures in a technical system: the entire process is well structured and each sub-step ends with a conclusion. Completed individual steps are not repeatable in the current session.

A relatively stringent regulation of the treatment steps is required because only semi-skilled or non-professional medical personnel will be working at the patient's side. A field test of the prototype under real conditions will show whether this step-by-step process is suitable for medical treatment and whether it will be accepted by the participating medical staff.

The doctor can unlock the required medical devices via the PMS at the PS site and provide the assistant with additional context-related support instructions. The assistant at the PS will be also be instructed by the doctor, in order to make no mistakes when carrying out the medical measurements.

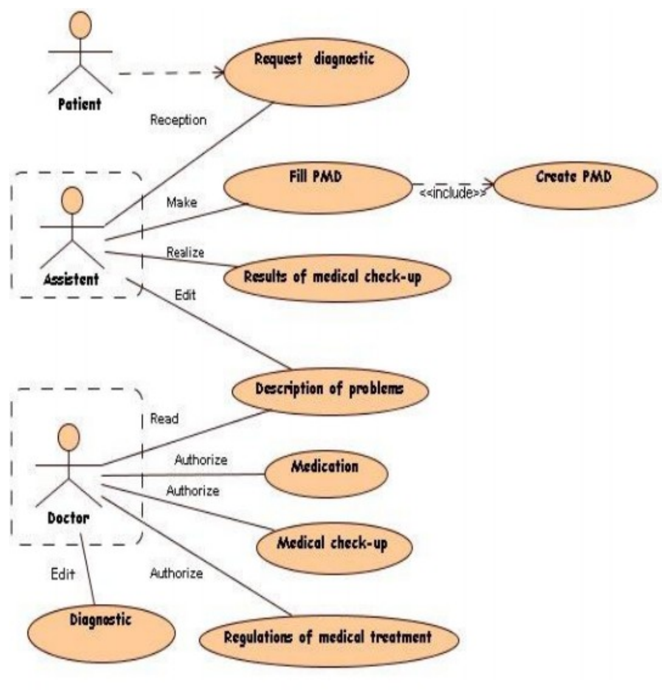



Figure 3. Graphic description of the asynchronous mode

The MDS consists mainly of the following components:

- *MD driver*: Implementation of the Corscience protocol interface to medical devices.
- *MD data server*: Recording and preprocessing of the measured medical data and implementation of a TCP/IP command interface.
- *MD proxy*: implementation of the TCP/IP command exchange between the MDS web page and MD data server.
- *MDS web pages*: The operation of each medical device is made via a web page with an MD device operator panel.

All MDS components are located on the PS and will be activated via the MDS web pages in the web browser of the PS. For this purpose on the PS is running a local web server which is used in order to start the MDS web pages. The PMS (loaded from the remote server TeleMedSys) and the MDS

 ISBN: 978-1-61208-235-6 

(loaded from the local web server in the PS) are connected in a frame website according the MashUp principle. Only by this way is possible for this web page to provide access to the peripheral components (medical devices) of the PS.

Figure 4 shows the frame website of the PS. The visual design of the PS website is not yet optimized. At the moment it is a design difference between the MDS (right grey part) and the PMS (left light-blue part).The reason is these two parts are implemented using different web technologies – MDS through pure HTML/JavaScript and PMS through Flash. The next version of the PS/DS will also have a unified user interface design.

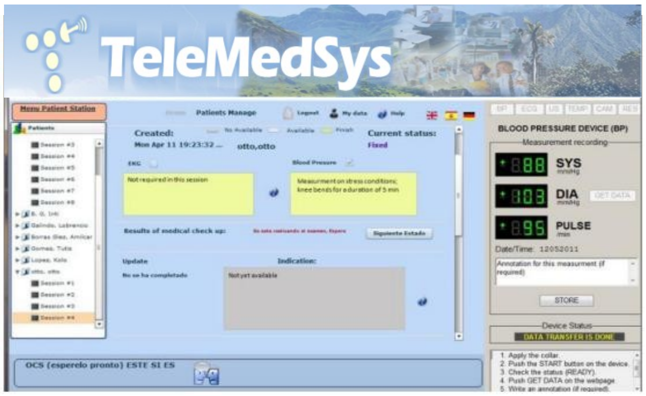



Figure 4. The frame website of the Patient Station in the prototype

In case of a limited data value volume, the MDS transmits the measuring data (e.g. blood pressure) in real-time to the PMS and stores the data in the database. The real time term is just for representing the immediacy of the information in the network between the PMS and MDS. In case of larger data volumes (e.g. ElectroCardioGram ECG) the data are saved locally in a measuring file first and subsequently transferred remotely in a corresponding file of the TeleMedSys server. In this case, only the name and location of the data file (not the data itself) are stored in the database.

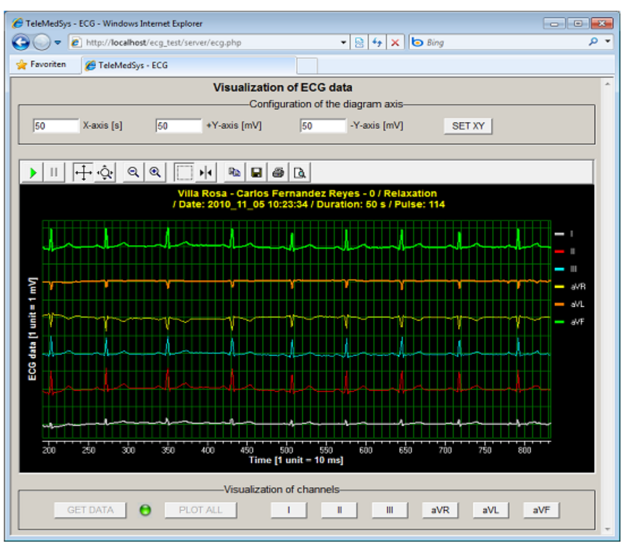



Figure 5. Analysis of a 6-channel ECG with a measuring time of 50 s in the Doctor Station’s web browser

## C. *Problems of data analysis*

The measuring data can be analyzed at the DS. Figure 5 shows an analysis example for the 6-channel ECG in the DS web browser.

The XY plot object (ActiveX object) from the Iocomp visualisation package is used for the graphical ECG data presentation [13]. Here, however, a problem arises for the visualization of long-term ECG measurement, which has not yet been optimally solved:

The recording of ECG data occurs at a clock rate of 100 Hz or 500 Hz. A normal stress ECG with e.g. 20 min duration and a resolution of 10 ms (100 Hz) results in 120,000 data values. The representation of these data in the XY plot object is in principle not a problem, but the plot object requires about 2 ms for the visualization of each data point in the XY object. Add to this the retrieval of data from the TeleMedSys server, which takes about 30 seconds. For the above example a delay of 4.5 min for the initial display of the data data thus follows. Once the measurement data in the data RAM of the XY plot object are stored, the analysis of data can be done quickly and in a very detailed way.

## D. *Design and Implementation*

The PMS realisation is based on HTML/Flash (Macromedia eXtensible Markup Language - MXML/ActionScript) in combination with Java for accessing a database. mySQL is used as database server and the design of the relational database was made according to the specification of business applications.

The MDS is developed with HTML, JavaScript and Java applets. The MD server is a Java application that is operated via a Bluetooth interface and controlled through TCP/IP commands by a Java applet (MD-Proxy, see Figure 2).

The technology selection for the development of PMS was based on capabilities offered to end users as well as developers by using Rich Internet Applications (RIA) such as Flex Builder 3.0; that was the tool selected for PMS implementation. The use of the Web 2.0 tool in a system like TeleMedSys; includes the optimization of time and resources. The applications development, with RIA, provides some advantages to the systems user and reduces information traffic on networks, because it is not necessary to reload all of the page but just that portion that needs to be updated.

The MDS uses a special Java applet with the Live Connect method for communication between the medical devices and JavaScript functions on the MDS webpages. Originally this Java applet was developed for the remote operation of industrial devices by TCP/IP commands

TeleMedSys is structured according to classic architecture for web systems: architecture of three levels supported in a Server – Client architecture. In order to ensure the usability of TeleMedSys, the architectural pattern Model View Controller (MVC) was selected. This pattern guarantees an effective information management; it separates the model of the user view layer, using a controller to intercede between these two levels.

 ISBN: 978-1-61208-235-6

To ensure the persistence of the PMS the binomial Hibernate/Spring framework is used. The main task of Hibernate is the Object-Relational Mapping (ORM). This makes it possible to store ordinary Java objects with attributes and methods in relational databases, and to create objects again from these records. Relations between objects are mapped to corresponding database relations. This guarantees that in the case of a connection abort or bad Internet connection no data are lost during treatment process, and the system can reset to the last treatment date.

The selection of the binomial also guarantees access to remote data transparently to the user; all procedures are performed using Remote Procedure Calls (RPC). Figure 6 shows the distribution of the PMS components in the TeleMedSys within the MVC pattern.

An Apache web server is running on the PS as a local web server. The PMS uses a Tomcat web server with a Spring Framework as Java application support.

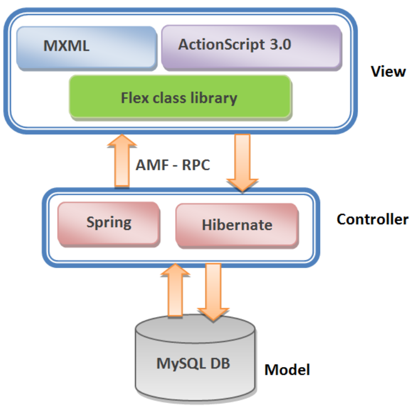



Figure 6. Architecture of the PMS

The DS will function as a TeleMedSys server simultaneously during a first field test of the prototype. Later, the TeleMedSys server will, however, operate on a dedicated computer; to provide the required safety and data security. For data transmission between the application and the server, the protocol implemented by Hibernate to access the data base was used.

TeleMedSys is a data transfer system; it must therefore ensure methods that ensure the integrity of these data. The system has three different kinds of users:

- Doctor user: Access to Patient Data Record (PDR), realization of diagnosis, records indications
- Assistant user: Access to PDR and MD, execute the indications of Doctor and has contact with the patient
- Administrator: Has the control of the system, adds, deletes or updates information about the rest of the users and the software configuration.

Each user can have access to the system using a password unique for each one. This password changes every time the user logs in the system. This password is encrypted using the MD5 algorithm.

For further versions of this system, we recommend using a data encryption algorithm in order to protect the patients' data, as well as the information of the diagnosis and the security of the system in general.

A database was designed in consideration of the data integrity. The final design of this database model reaches a 3rd normal form of normalization level. The design of the database takes into account the requirements of the system itself.

Figure 7 shows the final model design for the database

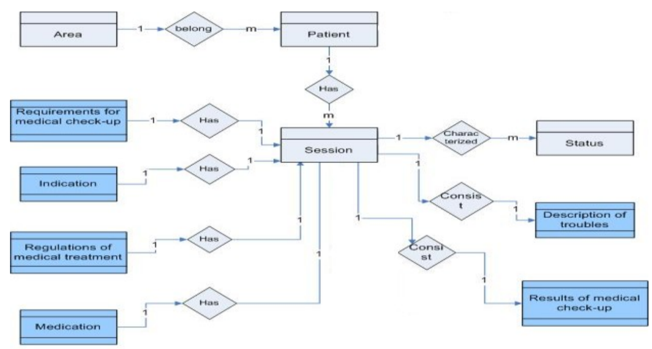



Figure 7. Database Entity Relationship Model

## IV. CONCLUSION AND OUTLOOK

The article describes a web-based telemedical system for use in rural areas. The combination of remote-based and local web applications enables direct remote access to medical devices through a web browser without technological gaps. The doctor station and N distributed patient stations are provided with a structured access to all patient data through a common web browser, enabling them to perform asynchronous treatment procedures.

Upon completion of the prototype (May 2011), a first field test in cooperation with the CENPIS of the Universidad de Oriente in Cuba was realized with real patients and the positive opinion of the doctors in July 2011. The test results showed that TeleMedSys represents a big step for rural health, not only because the population will benefit from better health conditions, but also from advances in the technology. This result will be integrated in a revised prototype, which will be deployed in one of the selected rural regions in Ecuador in 2012/13.

The logic distribution and architecture of the components of TeleMedSys, make it a simple and useful system that is to easy maintain and easily migrated to other media and technologies. TeleMedSys is easy to attach to mobile platforms. It is therefore conceivable to implement a system for remote diagnosis and telemedicine using technology and perhaps even mobile telephony.

In this case, the possibilities offered by these devices (cameras, sensors, microphones) will be useful for the next versions of TeleMedSys; this would make them practical, inexpensive and easy to use for all users. The proposal considers that only the PS uses the mobile device and that Android is considered as a platform for devices.

PS will have the following components: Mobile device – Smartphone and Web browser with HTML/Flash Player by

 ISBN: 978-1-61208-235-6 

default. The device should also include the software for MD and a local Web server. The Web server in PS can be an Android Web server (KSW [29] or I-Jetty [30]) or can be TomCat. The selection should be accord with the designer and also with the priorities of the TeleMedSys.

The implementation of this solution brings inconvenience to the users, which must be analyzed.

- The connection between the mobile device (Smartphone) and TeleMedSys server should be using Wi-Fi, which means data transfer will be slow and it is also possible to lose some information in the process. Of course, that depends on the quality of the Wi-Fi.
- The TeleMedSys server should be near the radio zone of the Wi-Fi.

To put this proposal into practice a number of actions must be taken first. For instance, the general security of the system must be ensured, because the mobile environment is very hostile and vulnerable to attacks. We recommend, once the first version of TeleMedSys is complete, that the necessary bases for the migration to mobile platform are created because of the advantages offered by it.

## REFERENCES

[1] *Zydacron*: Information about CareStation 140. – www.zydacron.com, 2009

[2] *Bellazzi, R., et al.*: Web-based telemedicine systems for home-care: technical issues and experiences. - Computer Methods and Programs in Biomedicine, 64 (2001), pp. 175–187

[3] *TMA Medical*: Erfolg für Telemedizin und E-health durch das Mobile Care Unit. – www.tma-medical.com, 2010

[4] *MEYTEC*: Ein neuesTelemedizinsystem für Echtzeitanwendungen in Rettungsfahrzeugen. – www.meytec.com, 2010

[5] *GS Elektromedizinische Geräte G. Stemple*: Information about corpuls systems. - www.corpuls.de, 2009

[6] *Williams, D. R., et al*: A Strategic Vision for Telemedicine and Medical Informatics in Space Flight. - Telemedicine Journal and e-Health, Vol 6, Nov 4, 2001, pp. 441-448

[7] *Centre Spatial Guyanais*: Telemedicine in French Guiana. – www.cnes-csg.fr, 2010

[8] *Cali, C.*: The Italian Military Medical Services and Telemedicine. – Proc. of the Telemedicin Workshop and 3rd Meeting of the SIMIHO new Experts technical Group, Rome, 7 - 8 July 2003

[9] *Melcer, T., et al.*: A Prospective Evaluation of ENT Telemedicine in Remote Military Populations Seeking Specialty Care. - Journal and e-Health, September 2002, 8(3), pp 301-311.

[10] *Logan, J. S.*: Telemedicine 101: Innovations in Rural Health Care Delivery Systems. – Proc. of Cooperative Solutions to Rural Health Care Problems: Emergency Medical Services, University of California, Davis School of Medicine, June, 1993

[11] *Blue print*: The Ohanyido Matrix: Development of a Telemedicine System for Nigeria. - http://ohanyido.tripod.com/id13.html, 2005, 2011

[12] *Pandian, P. S., et al*: Store and Forward Applications in Telemedicine for Wireless Ip Based Networks. - Journal of Networks, Vol. 2, No. 6, Dec. 2007, pp. 58 - 65

[13] *IOcomp*: Description of the iPlot object. – www.iocomp.com, 2010

[14] Estudio de Factibilidad de una Red de Telemedicina en la Red Pública del Atlántico. http://archivo.revistaesalud.com/index.php/revistaesalud/article/view/250/566, 2011

[15] *Martínez, Andrés:* Evaluación del impacto del uso de las tecnologías apropiadas de comunicación para el personal sanitario rural de países en desarrollo. - Tesis doctoral, Universidad Politécnica de Madrid, 2003. pp. 16-20.

[16] *Arado López, Annabel*: Propuesta para la implementación de la telemedicina como sistema a partir de la adecuación del soporte de comunicaciones de la red Infomed. - Tesis de Maestría, Departamento de Telemática, ISPJAE, 2007

[17] *J. Rottger, AM Irving, J. Broere, and B. Tranbmer:* Use of telecomunications in a rural emergency. - Brain surgery by fax!. Journal of Telemedicine and Telecare 3, pp. 59-60, 1997

[18] *D. Shanit, A. Cheng and RA Greenbaum*: Telecardiology: suporting the decision-making process in general practice. - Journal of Telemedicine and Telecare 2, 1996, pp. 7-13.

[19] *Treatment by telegraph (1917):* Excerpt from the obituary of John Joseph Holland (1876-1959). Journal of Telemedicine and Telecare 3, 1997, pp. 223.

[20] *F. Amenta, A. Dauri, and N. Rizzo:* Organization and activities of the International Radio Medical centre (CIRM). - Journal of Telemedicine and Telecare 2, 1996, pp. 125-131.

[21] *Informatización en Cuba*. Disponible en: http://www.cubaminrex.cu/Sociedadinformacion/Cifras 2005.

[22] Telemedicina y Bioingeniería. Bioingeniero Eduardo D. Lázaro. Disponible en: http://www.simiconsultora.com.ar, 2001

[23] *Revista Cubana de la Informática Medica:* Las Nuevas Tecnologías de la Información (NTIC) en la medicina: la Telemedicina en Cuba. - http://www.cecam.sld.cu/, 2011

[24] *RH Friedman, LE Kazis, A. Jette, MB Smith, J. Stollerman, J. Torgerson,and K. Carey*: A telecomunications system for monitorin and counseling patiens with hypertension. - American Journal of Hypertension 9, pp. 285-292, 1996.

[25] *D. Shanit, A. Cheng and RA Greenbaum:* Telecardiology: suporting the decision-making process in general practice. - Journal of Telemedicine and Telecare 2, pp. 7-13, 1996.

[26] Fundatel Blog de los viajes de los integrantes de Fundatel, fundación de telemedicina de Argentina, aplicando Telemedicina y eSalud por el mundo. Avaible in: http://fundatel.blogspot.de/

[27] *M. Mijares:* Limitations for the implementation of Telemedicine and Telehealth Projects: Based on the Ecuadorian experience. Avaible in: http://www.rev-conference.org/REV2008/archive/welcome-5.htm

[28] CorScience- Cardiovascular Innovations. Avaible in: http://www.corscience.de/

[29] kWS, Android Web Server http://www.androidpit.de/de/android/market/apps/app/org.xeustechnologies.android.kws/kWS-Android-Web-Server

[30] I-Jetty: webserver for android plataform: http://code.google.com/p/i-jet

# Ambient Traffic Services Based on the Use of Agile Warning and Notification System Radio-Help

Jan Skrbek, David Kubat, Tomas Zizka, Jiri Kviz, Klara Antlova
Department of Informatics
Faculty of Economics, Technical University of Liberec
Voroněžská 13, Liberec, Czech Republic
jan.skrbek@tul.cz, david.kubat@tul.cz, tomas.zizka@tul.cz, jiri.kviz@email.cz, klara.antlova@tul.cz

***Abstract*** **- The paper describes an efficient way on how to reduce the number of car accidents caused by collision with damaged vehicles. It aims at assessing the current situation and to identify weaknesses and influence of planned or proposed solutions. The agile "System for Automated Forewarning of Vehicle Crashes" is introduced. The input data are collected from eCall system. Distribution of warning information is provided through Radio-Help broadcasting that ensured distribution of warning messages to drivers arriving to the place of an accident. With respect to the existing warning systems and its weaknesses, the ambient traffic warning service would allow us to get proper information on time during any kind of road problems.**

***Keywords-agile; eCall; crashes; warning; information; Radio-Help***

## I. INTRODUCTION

One of major problems of the contemporary world is a steady increase in numbers of vehicles and the associated increased load of roads and motorways. Large society of car drivers demands such kind of ambient traffic services that could efficiently reduce the risk of serious traffic accidents. Among most common causes of serious traffic accidents associated with injuries and in adverse event with deaths of their participants are accidents caused by collision with damaged vehicles. Timely distribution of relevant information is a key to reduce the number of traffic accidents and economic and human losses in such situations. Unfortunately existing warning systems are not capable to cover such requirements.

As a case study model a recent event was used. On November 5th, 2011, around 6 p.m., a huge crash involving 34 vehicles on a highway in Southwestern England killed at least seven people and injured 51 [3]. The crash involved explosions, and cars and tractor trailers burning "literally to the ground." Police expected the death toll to rise and they feared they may find more bodies in the wreckage. It was not immediately clear what caused the collision on the M5 highway, but police said foggy conditions and wet road surfaces were partly to blame. The affected section of the busy highway was closed in both directions as police removed all vehicles for forensic examination.

Chain car crashes are a typical example of situations with high demand for crisis communication. Crisis communication can be broadly defined as the collection, processing, and dissemination of information required to address a crisis situation [2][11]. Let us consider these situations as our case study model and compare their real progress with the one that would have happened if eCall and Radio-Help systems were implemented.

Currently, the information about a traffic accident is reported verbally to the emergency operations centres via mobile phones, either by those involved in accidents or their witnesses (Fig.1). However, this is associated with problems when attempting to explain the given situation and determining adequate intervention (the exact position and direction of the vehicle, the scope of damage, elimination of repeated reports of the same accident, etc.). The speed of intervention is a key factor for its success, whereby any possible delays influence negatively the outcome of the entire rescue operation and the health status of injured persons.

Currently, there are several projects [1][4][10] in various stages of development aiming to solve the current traffic problems in order to reduce damage to property and to protect health and lives of road users. The core of many of them is geographical positioning of traffic accidents. This data can be used for warning and noticing drivers prior the place of traffic accident.

Next section brings information of current solutions and identifies its weaknesses. Section III provides information related to proposed solution. Last section brings conclusion and our plan for future work.

## II. DESCRIPTION OF SELECTED TELEMATICS SYSTEMS

### A. *Variable information boards*

In Czech Republic, information displayed on the boards on motorways is received from the Unified Traffic Information System, a joint project of the Ministry of Transport, Directorate of Roads and Highways and several other bodies and organizations.

Currently, there are about one hundred of these variable information boards installed on motorways in the Czech Republic, representing coverage of approximately one board per 20 kilometers of a highway or a motorway [5]. There are either no or very few smart boards on other roads and streets.

For example, in extreme traffic conditions during a normal working day an average number of 1,400 cars per hour passes the 96 kilometers of the D1 motorway. Delayed distribution of information in a matter of minutes, which is caused by time required for the processing and publishing of this information, brings danger for many motorists who can never receive information about the event in front of them via the variable information boards.

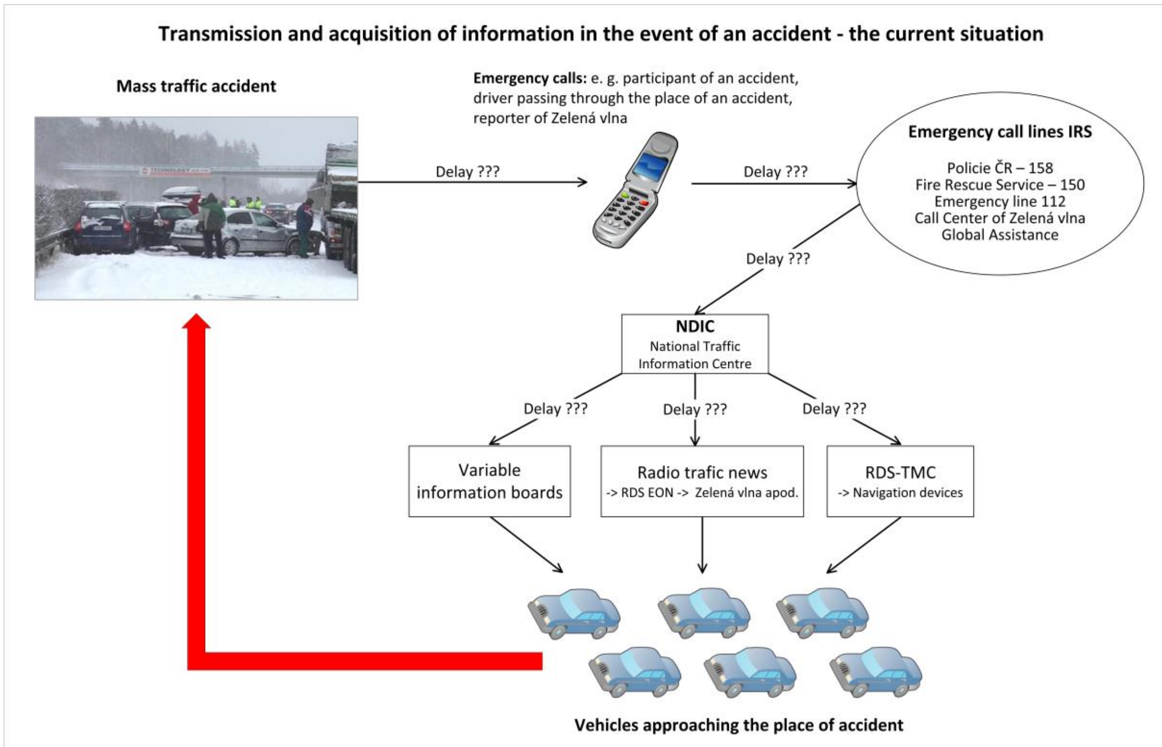



Figure 1. Transmission and acquisition of information in the event of an accident (Source: Authors)

*B. RDS-TMC*

RDS-TMC (Radio Data System - Traffic Message Channel) is a service that provides the drivers with traffic and travel information before and during their journey. This service integrates all relevant information and gives the driver a possibility to optimise the journey. The aim of the RDS-TMC is to provide traffic information within the FM broadcast band using RDS technology. Information is coded using an independent ALERT-C protocol and later on transmitted to the users as a silent part of FM broadcasting and further processed by the navigation device. According to national and international studies the main system benefits encompass significant improvement in traffic continuity and lower environmental impacts.

The disadvantage of this system is that a warning symbol appears in case a traffic problem occurs anywhere on the preselected route. For more information, the driver must manipulate the navigation device, which requires his attention. In addition, if there are further problems occurring on the given route, the warning icon remains unchanged despite the possibility that this newer traffic incident may have occurred in a location which is even closer in route than the originally reported traffic problem.

*C. eCall (Emergency Call System)*

Project co-funded by the European Union aims to create a system that enables automated reporting on accidents to the European-wide emergency line 112, including accurate information about its location. When the eCall device installed in a car detects an accident by means of sensors, it automatically sends a message to the nearest emergency centre, indicating the exact geographical location of the accident as well as other data. This system can be activated either manually by pressing a button on the dashboard by the vehicle passengers or automatically by the vehicle sensors triggered during an accident. After the system is activated, a connection with the nearest emergency call centre (PSAP) is established transmitting both sound and data flows. The sound connection enables vehicle passengers to communicate with professionally trained call operators while at the same time data channels are used to transmit data messages (MSD) to these same operators. Each message contains 43 details about the accident; such as time, exact location, car identification, eCall system status (whether the eCall was activated manually or automatically) and information about possible service providers. Based on this information, the operator will liaise with the integrated emergency services to direct them to the exact accident

 ISBN: 978-1-61208-235-6

location, as well as provide them with an exact description of the accident's severity and the number of injured [12].

Manual use of the system can be useful when we witness a traffic accident [4]. eCall systems should be installed in all new cars, at the latest, by 2015 and possibly also installed in older cars.

Although this system brings a clear improvement of the current situation in terms of saving lives and providing quick health care during accidents, it does not provide a solution for distributing information about the accident to the drivers approaching the place of accident, i.e., who are potentially in danger. When using existing information channels, the acquired accident data could be made available in some 5-10 minutes via motorway information boards, RDS-TMC messaging and radio travel news. However, each of these distribution channels has specific limitations and based on current traffic density, the above-mentioned reporting times are clearly insufficient.

### *D. Smart Road Restraint Systems*

The project aims - in addition to addressing timely reporting of accidents – at eliminating loss of life and property through timely preventive distribution of warning information. The proposed system obtains information about the current situation using existing visual and sensory infrastructures (highway camera system, radar system and weather condition monitors) and distribute such information to drivers. It also seeks to find opportunities for new materials to decrease safety hazards (such as better energy absorption through deformation zones of transport). This project is one of three priorities of the EU on the issue of transport in 2020 and is also co-financed from EU funds [10].

## III. AGILE SYSTEM FOR AUTOMATED FOREWARNING OF VEHICLE CRASHES

Agile approach is based on new practices and techniques that make product development more cyclical and incremental. It relies on lean governance (management) as opposed to more traditional techniques that rely on heavyweight governance. Agility is also about empowering the team and getting closer to what the customer wants. In place of rigorous upfront planning and the phase-based process, it offers a dynamic, iterative build-and-test cycle, where a change is handled well. One of Agile's hallmark features is that it drives the decision-making process lower in an organization, making that organization more responsive and adaptive.

For better and particularly early distribution of warning information, a system called "System for Automated Forewarning of Vehicle Crashes" (the System), which has a data connection to the receiver systems-vehicle emergency call (e.g., eCall) could efficiently help. The principle consists in full automation of generation and transmission of all relevant information about the accident to vehicles moving in its vicinity. The process of warning is initiated by the crashed vehicle, which generates data about the accident using eCall immediately after the collision happens together with the exact location of the accident. Information is received by the central office of the System which immediately generates data and/or voice information about the incident, including the positional code of the accident. Data are immediately broadcasted via radio session to car receivers [1].

System receivers (mobile phones, car radios, navigation devices) must be equipped with GPS receiver and a positional code comparator of an accident positional data generated by the positioning system receiver. If the comparator evaluates that the position code of an accident coincides with the position code of the receiver and vehicle movement will be evaluated as being directed to the scene of the accident, it will be forced to activate the data reception and/or voice session. In practice, we may be able to automatically inform road users according to their current position and direction of the danger which is coming, almost immediately. The system is functional and usable for any kind of road or street, thus, not only on highways and motorways.

The System uses HD radio broadcast technology or generally any digital radio broadcasting system, supplemented by determining the position through GPS or Galileo. If we consider data acquisition for warning from eCall, in the event of a major expansion an addressable warning system that would significantly limit the creation of public transport accidents could be a very effective. Transfer of information in the case of using "System for Automated Forewarning of Vehicle Crashes" is shown in Fig. 2, where solid lines show the flow of information the driver will receive in real-time or with minimal delay.

Detailed principle of radio broadcasting of warning information is described in [7] under the working title RADIO-H (Radio-Help). It is based on simultaneous application of analogue broadcasting technology with superposition of digital content (HD RADIO or DRM) or full-digital broadcasts with the possibility of defining the positional coordinates via GPS [7][8][9]. HD Radio technology company iBiquity Digital Corporation was selected in 2002 in the U.S. as a key technology for the digitization of radio broadcasting. Currently, this technology carries a large percentage of U.S. radio stations.

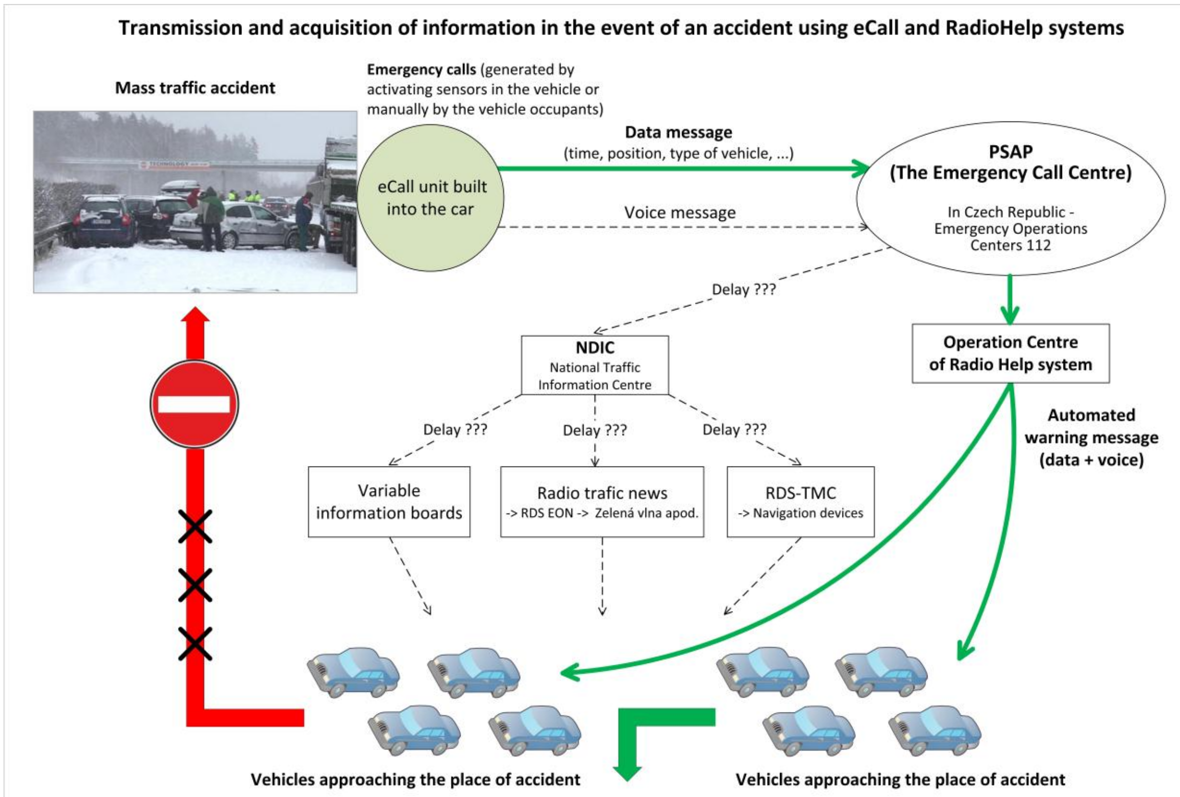



Figure 2. Transmission and acquisition of information in the event of an accident with the use of eCall and Radio-Help (Source: Authors)

HD Radio technology uses the principle of superposition of the digital signal to analogue signal. „Today, there are in the USA more than 2,100 stations with HD Radio Broadcast Technology… HD Radio™ Technology is available or has been announced to be available on 28 different brands of new vehicles (Acura, Audi, Bentley, BMW, Buick, Cadillac, Chevrolet, Ford, GMC, Hyundai, Jaguar, Kia, Land Rover, Lincoln, Mazda, Mercedes-Benz, Mini, Porsche, Ram, Rolls Royce, Scion, SRT, Subaru, Tesla, Toyota, Volkswagen, Volvo). These announcements represent 148 different vehicles with HD Radio Technology, a total of 74 which will come with HD Radio Technology as a standard feature during 2012 [13].

The transmitted relation of Radio-Help uses positional codes for identifying areas of compulsory income, i.e., where the broadcast is directed. The receiver in the area is maintained in a standby mode and captured broadcast on fixed rate compares its position according to GPS coordinates with areas included in the broadcast. If there is compliance, it activates forced broadcast reception session. After the broadcasting code ends, the receiver goes into standby mode again. Subscribers of Radio-Help that are outside the defined zone will not be disturbed by warning broadcast sessions.

This principle implies that it is possible to transmit separate sessions to more areas simultaneously. Long wave radio transmitters, which with new higher quality broadcasting channels gradually lose their utility, could be used for the broadcast. In such a case, it would suffice to cover the whole Czech Republic just by one central long wave radio sender.

Due to the development of IT where circuits for terrestrial broadcasting and positioning GPS are now equipped with most new mobile phones, it should not be technically demanding to use it for these purposes.

There are often doubts about privacy with respect to the positional information. It needs to be said, that the distribution of such information is done by the eCall system itself. From that perspective, as the eCall system has been already approved, the discussion of privacy would be redundant.

The System for Automated Forewarning of Vehicle Crashes is a novelty idea based on some previous work. Compared with other systems considering traffic (car accidents and similar issues) – only this approach allows fully automated warning limiting the influence of human factor. This fact provides shorter response time therefore better functionality.

## IV. CONCLUSION AND FUTURE WORK

While the RDS-TMC has long been in operation and eCall should be installed in new cars from 2015, a project of Smart Road Restraint Systems is still under development. Automatic emergency call system eCall is designed especially for accelerating action of rescuers and other components of IRS (Integrated Rescue System). The availability of accurate information about the accident, particularly the place of it and some additional information like a type of car or extent of the damage, without any significant delays will undoubtedly be very beneficial. It follows that the introduction of eCall may help reduce human losses and reduce the consequences of accidents by early intervention emergency services.

The prevention of subsequent accidents may at least have an equivalent effect. If the data from eCall is also used for early warning of other potential participants in an accident, it can lead to significant lowering of life, health and economic losses that could follow in the event of traffic accidents. Combining eCall together with the use of “System for Automated Forewarning of Vehicle Crashes” can ensure real-lime distribution of warning messages to drivers coming to the accident. Drivers should have relevant warning information in time in case of approaching the accident, so they could respond in time. The described combination can provide direct transfer of relevant information with minimum delay. The system is functional and usable for any kind of road or street. In addition, information is sent only to specifically defined geographical area - for example, only to drivers of vehicles that are closer to the accident than 15 km and also go towards this accident. The combination of the System with car navigations could deliver the necessary information only to drivers approaching the place of the car accident. The System has also an ambition to be an ambient service for improving the first aid to injured persons.

As a next step proving the quality of the system there have to be done a modeling of the situation with the System implemented. It should provide reliable data needed for a quality analysis.

## ACKNOWLEDGMENT

The presentation of the paper was supported by the project No. CZ.1.07/2.2.00/28.0327 Innovation and support of doctoral study program (INDOP), financed from EU and the Czech Republic funds. Other support was SGS project Agile Approach of Information Support by Solving Non-Standard Situations.

## REFERENCES

[1] M. Brunclík and J. Skrbek, 2010. “Systém automatizovaného kritického varování před místem dopravní nehody” [online], [retrieved: June, 2012], Available from WWW: <http://spisy.upv.cz/Applications/2010/PPVCZ2010_0415A3.pdf>.

[2] W. T. Coombs and S. J. Holanday, 2010 (ed.). “The Handbook of Crisis Communications“, Willey-Blackwell, Chichester, 2010, p. 20, ISBN 978-1-4051-9441-9.

[3] Conservative Refocus News, 2011. “Fiery Chain-Reaction Accident Results In 34 vehicle Pile-Up in United Kingdom“, [retrieved: May, 2012], Available from WWW: <http://www.conservativerefocus.com/blog5.php/2011/11/05/fiery-chain-reaction-accident-results-in-34-vehicle-pile-up-in-united-kingdom-at-least-7-dead-scores-injured-video>.

[4] European Commision, 2010. “eCall – saving lives through in-vehicle communication technology“, [retrieved: May, 2012]. Available from WWW: <http://ec.europa.eu/information_society/doc/factsheets/049-ecall_july10_en.pdf>.

[5] Ředitelství silnic a dálnic ČR, 2012. “Proměnné dopravní značky (PDZ) a zařízení pro provozní informace (ZPI) “, [retrieved: June, 2012], Available from WWW: <http://portal.dopravniinfo.cz/promenne-dopravni-znacky-a-zarizeni-pro-provozni-informace>.

[6] SafetyNet, 2009. “eSafety”, [retrieved:May, 2012]. Available from:<http://ec.europa.eu/transport/road_safety/specialist/knowledge/pdf/esafety.pdf>.

[7] J. Skrbek, 2011. “Advanced Ways and Means of Civilian Notification in Crisis Situations“. In A. KOCOUREK (ed.). Proceedings of the 10th International conference: Liberec Economic Forum 2011. 1st edd. Liberec: Technical University of Liberec, 2011. pp. 419–426. ISBN 978-80-7372-755-0.

[8] J. Skrbek, 2009. “New Possibilities of Information Services for Special Situations“, in 17 – th Interdisciplinary Information Management Talks, Proceedings (IDIMT-2009), pp. 123–130, Trauler Verlag, Linz, ISBN 978-3-85499-624-8.

[9] J. Skrbek, 2011. “Radio-Help as a Smart Early Warning and Notification System“. In Proceedings of the 55th Annual Meeting of the International Society for the Systems Sciences, 14 p, University of Hull Business School, UK, 2011, http://journals.isss.org/index.php/proceedings55th/issue/view/11, ISSN 1999-6918.

[10] SMART Road Restraint Systems, 2010. Innovative concepts for SMART Road Restraint Systems to provide greater safety for vulnerable road users, [retrieved: May, 2012]. Available from: <http://smartrrs.unizar.es/up_files/file/WORKSHOP/folleto_smartrrs3.pdf>.

[11] K. J. Tierney, M. K. Lindell, and R. W. Perry, (2001) (ed.) “Facing the Unexpected“, Joseph Henry Press, Washington, 2001, p. 13, ISBN 0-309-06999-8.

[12] J. Vávra, 2010. “Systém eCall v konfrontaci s alternativními systémy nouzového varování“, DP-EF-KPE-2010-105, EF-TUL, Liberec.

[13] iBiquity Digital, “What is HD Radio Broadcasting?”, [retrieved: May, 2012], Available from WWW: <http://www.ibiquity.com/hd_radio>.

# An Ambient Assisted Living Framework Supporting Personalization Based on Ontologies

Christophe Jacquet, Ahmed Mohamed
*SUPELEC Systems Sciences*
*Gif-sur-Yvette, France*
*christophe.jacquet@supelec.fr*
*ahmed.mohamed@supelec.fr*

Marie Mateos, Bruno Jean-Bart
*TRIALOG*
*Paris, France*
*marie.mateos@trialog.com*
*bruno.jean-bart@trialog.com*

Pierre Bretault, Indrawati Schnepp
*ATEGO*
*Issy-les-Moulineaux, France*
*pierrebretault@neuf.fr*
*indrawati.schnepp@atego.com*

Yacine Bellik
*LIMSI-CNRS*
*Orsay, France*
*yacine.bellik@limsi.fr*

***Abstract*—As the population in many countries is steadily aging, allowing elderly people to stay longer at home is a growing concern. Ambient Assisted Living (AAL) proposes new techniques to help people remain autonomous, based on ambient intelligence. We present an ontology-based framework in which ontologies enable the expression of users' preferences in order to personalize the system behavior. They are also used for the discovery and interconnection of devices, the storage and retrieval of collected data and the transmission of actions. This way, the behavior of the system may be expressed using high-level logic rules. Another important contribution is the addition of a diagnosis service that monitors the run-time behavior of the AAL system only by using sensors discovered *opportunistically* at run-time and knowledge about physical laws, not pre-defined control loops. Finally, this paper describes an actual implementation, with precise technological details, in order to prove the feasibility of the technical choices, and provide implementation ideas for future projects.**

***Keywords*-Ambient Assisted Living (*AAL*); ambient intelligence; ontologies; diagnosis; reasoning.**

## I. Introduction

Due to the demographic change towards an aging population, society must find ways to assist elderly people to stay active at home longer. While currently this support is mainly provided by human caregivers, technology will play a more and more important role both for elderly persons and caregivers. In Europe a roadmap has been defined in the last years called Ambient Assisting Living (AAL) [1]. The business context of AAL is rich in terms of technology (from tele-health systems to robotics) but also in terms of stakeholders (from service providers to policy makers, including core technology or platform developers).

The work presented here has been carried out within the CBDP project (Context-Based Digital Personality) [2], which aims at creating a framework for building various kinds of ambient-intelligent applications, based around the concept of Digital Personality for representing the preferences of users. Aside from AAL, several application domains were considered, such as digital TV guides or assistants for workers at a construction site. Therefore the CBDP framework addresses a wide variety of requirements. In this paper however, we focus exclusively on the parts of the CBDP framework relevant to AAL.

Our approach is entirely based on ontologies. Not only are ontologies used to capture domain knowledge, but more importantly they serve as the runtime mechanism that allows the interconnection of devices, the exchange of data and the execution of actions. Moreover, by examining the sequence of requested actions and observed sensor values found in the ontology, a diagnosis process is able to monitor the run-time behavior of the system and to detect unexpected patterns.

The ontology is presented in Section II. Section III describes the CBDP framework and gives implementation details; Section IV focuses on diagnosis. Section V describes a typical AAL use case, and goes through its complete realization. Section VI introduces some related work, and compares our approach with published results. Finally Section VII gives directions for future work.

## II. An Ontology for AAL Applications

CBDP is built around an ontology: this section justifies this choice and describes the ontology used.

### A. Why use ontologies?

AAL applications are trans-disciplinary by essence (for instance, they can mix automatic control with modeling of user behavior), therefore the ability to reuse knowledge and integrate several knowledge domains is particularly important for them. Furthermore, the field of AAL is very open and changing, so it is not possible to base an AAL platform on a fixed set of features, on a fixed set of data models: extensibility is key. In addition, an AAL environment may require the interoperation of software and hardware devices from a variety of suppliers: there must be a standard way of exchanging knowledge.

Ontologies are well-adapted to all these needs [3]: an ontology framework provides a standard infrastructure for sharing knowledge. In addition, semantic relationships such as equivalence may be expressed between various knowledge

 ISBN: 978-1-61208-235-6 

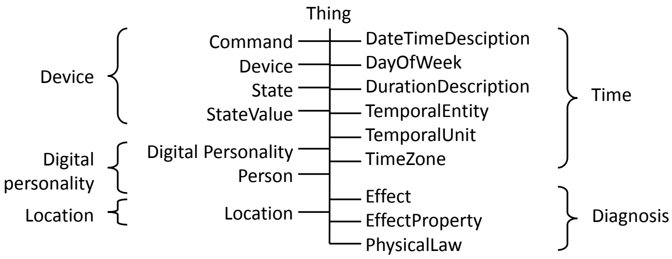



Figure 1. First level of the CBDP ontology.

sources, thus permitting the easy integration of several sources or domains. In addition, one can easily extend an ontology to take into account new applications or new devices. For these reasons, leading AAL projects such as OASIS (Open architecture for Accessible Services Integration and Standardisation) have put a strong emphasis on ontologies [4]. Being oriented toward personalization, CBDP explicitly introduces an ontology module for modeling the "Digital Personality" of the user.

### B. *Ontology used by CBDP*

The ontology defined for CBDP is built around the OWL language [5], which is based on the Resource Description Framework (RDF). RDF represents knowledge as a set of *triples* or *statements* of the form {subject, predicate, object}. It models different interrelated domains in a modular way, so as to enable its easy adaptation to new applications. In order to put into practice the aforementioned notion of reusability, two of the domains are based on existing ontologies. Figure 1 depicts the first level of the ontology; the main domains are as follows:

- *Device:* this part is based on the DogOnt [6] ontology that has been simplified for our purpose, while keeping the modeling axes (typology, functionality and state).
- *Digital Personality:* a class *Person* allows the representation of a human being, and a *Digital Personality* stores the person's preferences in order to personalize the services offered to him/her.
- *Location:* a location model is required because most of the services offered in the AAL domain must know the position of the user (in/out the house, in the bedroom/in the kitchen, etc.) and of the devices (sensors and actuators).
- *Time:* we import W3C's existing Time Ontology [7] without any change.
- *Diagnosis:* we introduce the concept of physical effect (see Section IV below), to compute the expected result of the action of an actuator onto a sensor.

The ontology is loosely coupled with the framework, so to a great extent it may be changed without affecting it. However, the basic feature of sending commands to actuators rely on specific *core classes and properties* that may not be changed: this part is depicted on Figure 2.

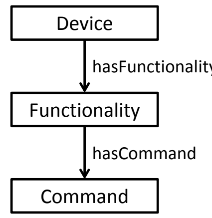



Figure 2. Ontology classes required for proper operation of the framework.

## III. CBDP FRAMEWORK

This section describes the CBDP framework, and how it can be used to build AAL applications.

### A. *Architecture*

The main goal of the CBDP Framework is to dynamically handle ontology data and initiate actions when specified conditions in the ontology are achieved. CBDP is written in Java; it is based on OSGi (Open Services Gateway initiative framework) [8], which allows one to build applications flexibly by combining *bundles*. In CBDP an application is composed of CBDP's core bundles (the Context Reasoner and the Sensor/Actuator Layer, described in Sections III-B and III-C, respectively) and application-specific bundles (see Figure 3). In our case:

- AAL-specific application bundle: contains the rules that define the intended application behavior, meant to assist the user according to his/her needs.
- Zigbee Driver bundle: allows the exchange of data between the physical devices (connected via a wireless Zigbee network) and the CBDP Framework.

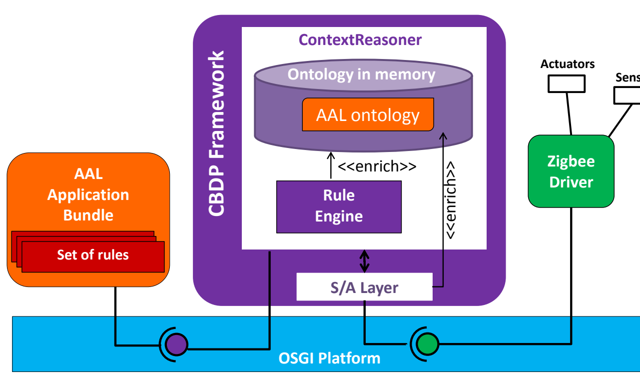



Figure 3. Architecture of the CBDP framework.

### B. Context Reasoner and Rules

The Context Reasoner is in charge of managing the information coming from external components (AAL Application or Zigbee Driver) by structuring them according to the AAL ontology. Therefore, it provides methods to add new information, retrieve stored information, and perform queries about that information. Manipulation of the ontology is done using the Jena library [9].

Another feature of the Context Reasoner is its rule engine. Its purpose is to perform actions to help the user and facilitate common tasks, based on a set of application-specific rules (hence the rules are provided by the AAL Application bundle). The rules are Horn clauses [10]: a rule is composed of premises that determine the situations in which the rule applies, and a conclusion, that basically adds a new "fact" into the ontology, such as a new property value. An example of such a rule is given in Section V-A below. Rules are applied by Jena's basic reasoning engine, using forward chaining.

For performance reasons, the rule engine does not apply all rules at each instant. The rules are applied only when a change in the ontology matches a *filter* (i.e., happens in a specific part of the class hierarchy). The filters are application-specific; here they are defined by the AAL Application bundle. At first one may use a "catch-all" filter; performance can be improved later by refining the filters.

### C. Sensor/Actuator Layer

The Sensor/Actuator layer (S/A layer) connects the sensors and actuators to the ontology. The communication is two-way:

- Sensor data (sent through Zigbee) is stored in the ontology. This allows one to perform semantic queries and semantic reasoning over sensor data.
- A command request inserted in the ontology (using a property called *hasCommand*) triggers the actual emission of a command to the actuator.

The module responsible for connecting the sensors to the Context Reasoner is based on the use of a specific OSGi service called *EventAdmin*. A communication protocol through OSGi events has been defined in order to allow the communication between the drivers and the S/A layer. Section III-D describes this protocol.

### D. Communication between sensors/actuators and the ontology

This section deals with the protocol used to exchange ontology knowledge using OSGi events. An event is composed of a *topic* and of a list of *properties* ({propertyName; propertyValue} pairs). We have defined two kinds of events: 1) to report sensor data, 2) to send commands to actuators. For both kinds of events, the OSGi *topic* string is built according to the pattern `CBDP/AAL/`*`deviceClass`*. `CBDP` and `AAL` are invariant: they reference the general framework and our application-specific ontology; *`deviceClass`* is the name of the sensor class that sends data, or actuator class that is to receive data. The remainder of this section gives details on the actual *contents* (list of properties) of the events in both cases.

*Sensor information*

| | |
|---|---|
| instance.id | URI identifying the sensor (String) |

*1) When referencing a dataProperty present in the ontology*

| | |
|---|---|
| data.property<br>data.property.value | Name of the "simple data" property (String)<br>Value (depends on property: Boolean, Double, Integer, String...) |

*2) When referencing an objectProperty present in the ontology*

| | |
|---|---|
| object.property<br>object.property.range | Name of the "object" property (String)<br>Name of the class referenced by the property (String) |

Table I
OSGi PROPERTY NAMES USED TO SPECIFY OWL TREES.

*1) Reporting sensor data:* When sensor data is reported, a sub-graph (actually a tree) must be created in the ontology. An edge in this tree may be of two kinds: connecting an object to a simple value such as an number ("dataProperty"), or connecting an object to another object ("objectProperty"). A convention using OSGi's properties allows us to completely describe the tree. At each node in the tree to be created, a set of datatype and object properties may be specified. Each edge of the tree is numbered using a simple convention: from the top of the tree, each time an edge is followed, a dot and the index of the edge under its parent node are appended to the OSGi property name (see the examples on Figure 4). This permits the description of each edge and each node to be created. The basic property names (without trailing dots) are given in Table I, and a complete example is given on Figure 4. It represents an event stating that the light level is 500 lux in the kitchen at the date {Calendar value}.

*2) Sending actuator commands:* Sending a command to an actuator is done using the following convention: a new statement must be added in the ontology, with a relation named "hasCommand" (see Figure 2 above). Such a statement may be added by a reasoning rule, or by application code calling the context manager.

When the S/A layer detects a new "hasCommand" statement, it serializes the corresponding sub-tree of the ontology graph into an OSGi event (using the same convention as above) and sends it to the driver of the target actuator.

### E. Deployment

The OSGi implementation used by CBDP is Apache Felix. The use of Java and OSGi permits to deploy the framework on a variety of platforms. We have conducted tests on desktop PCs (under Windows and MacOS) and on embedded systems (on a set-top-box running Linux and on Aonix Perc) [11]. Perc is a Java virtual machine

 ISBN: 978-1-61208-235-6 

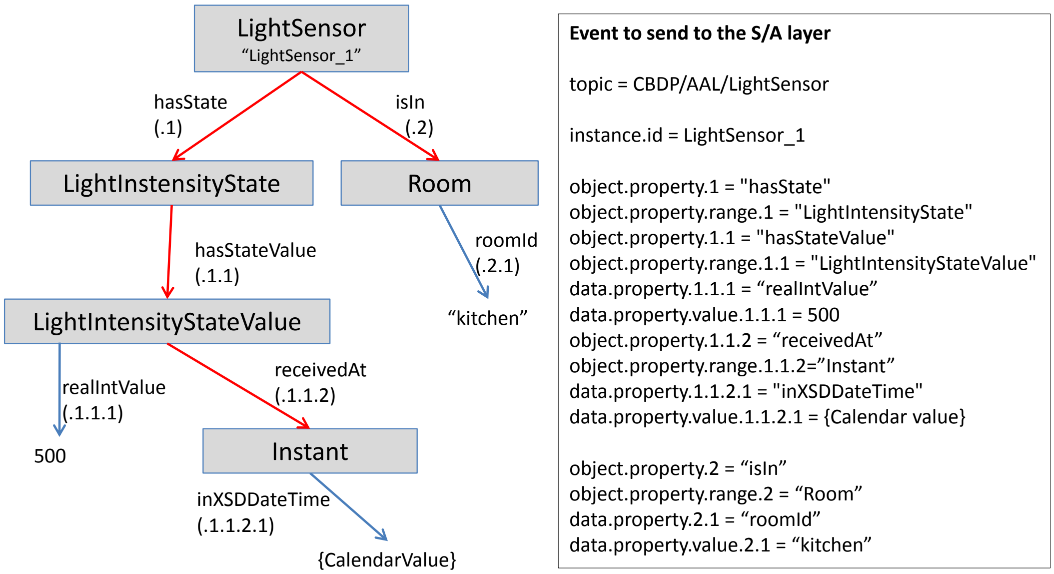



Figure 4. Example of an event containing sensor data.

for embedded systems that can be deployed on resource-constrained targets while providing real-time and safety guaranties. This demonstrates the adequacy of CBDP for its target applications, user assistance in ambient environments, i.e., in non computer-centric settings.

## IV. Diagnosis

Ultimately, the goal of any AAL application is to activate some actuators, based on data provided by some sensors. However, sensors and actuators may suffer failures. Therefore the system should check autonomously whether the intended actions are performed correctly.

### A. Rationale

In software, mechanisms such as exceptions and error codes report whether a procedure executes successfully or not. Likewise, an actuator can provide a return code, but generally this reflects only the way the orders are transmitted to the actuator, not their actual execution. For instance, when the system activates a light bulb, it receives an acknowledgement that confirms the switch-on of the electrical circuit, but this does not necessarily mean that the bulb is really on (the bulb may be damaged for instance). To address the issue, control theory could allow one to pre-determine closed control loops using designated sensors. However, the particularity of ambient systems is that physical resources, mainly sensors and actuators, are not necessarily known at design time, but are dynamically discovered at run-time, so such control loops cannot be pre-determined.

Therefore, a reliable AAL application needs a way to assess at run-time the status of its sensors and actuators. We propose an approach in which the system relies only sensors already available, thereby not requiring the addition of specific devices for diagnosis purposes. The sensors that may be used to perform diagnosis are discovered at run-time. When a sensor measures a physical parameter, the system may deduce sensor/actuator "health" status by comparing actual values with *expected* sensor values.

To achieve this, we propose a diagnosis framework in which the characteristics of actuators and sensors, as well as the *physical effects* involved, are precisely described. The following paragraphs provide a short summary of our approach; refer to [12] for more details.

### B. Modeling physical effects

*Effects* are modeled in order to simulate the physical consequences of actions in an ambient environment. Each effect is characterized by a set of properties: some *define* the effect (at the source actuator, e.g., the light intensity *emitted by a light bulb*), some are *observable* by a sensor (e.g., the light intensity *received by a light sensor*).

Depending on the application's needs, an effect can be defined at various levels of granularity. For instance, the light emitted by a light bulb could be modeled either using classical laws of physics for light propagation, or using a simple boolean law ("if a light bulb is on in a room then the light sensors that are in that room should detect light").

### C. Using effects for linking actuators to sensors

As ambient systems are highly dynamic, one cannot explicitly link related sensors and actuators. The concept of effect allows for easy decoupling of devices, as illustrated by Figure 5. An actuator class is linked to the effects it may potentially produce. Similarly a sensor class is linked to at least an effect property. At a generic level, there is a link between a given effect (e.g., emission of light) and the corresponding detectable properties (e.g., light intensity) through the *hasProperty* relation.

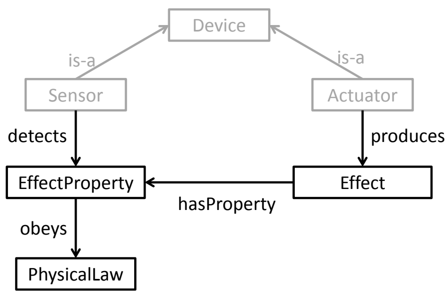



Figure 5. The diagnosis framework adds to the general-purpose ontology (gray part at the top) a few concepts to describe the effects and their detectable properties.

Knowing the effects produced by any actuator in the system, and knowing the effect properties sensed by any sensor in the system, it is therefore possible to determine and update at run-time the links between actual sensors and actuators. Hence it is possible to compute the expected readings of the sensors using physical laws. Once the expected results have been determined, the system checks if they are consistent with the actual readings.

## V. Experimentation

This section introduces a complete AAL scenario in which the CBDP framework is able to automate tasks, it shows how diagnosis is performed, and it describes the experiments.

### A. Use case: automatic light switch

We propose the following experimental scenario. It takes place in a bedroom with a controlled lamp, a light sensor and a presence sensor: "if the ambient light level is under a threshold (specified in the Digital Personality of the user) and if the user is present in the room, then the light must be turned on". Although simple, this scenario demonstrates all the aspects of the system: sensor data gathering, reasoning, command of actuators and diagnosis.

Let us suppose that the light level in the room is 80 lux and then the user comes in. His Digital Personality states that he wants the lamp to be on when the light level is under 100 lux. The system takes the following steps:

1) The current light level (80 lux) has already been detected and updated in the ontology. When the user enters the bedroom, the presence sensor sends a notification to the driver through the Zigbee network. The driver sends then an event to the framework and the ontology is updated accordingly.
2) The framework detects that the value of a PresenceSensor has changed in the ontology, so the following rule must be evaluated (cf. III-B):

```
IF a LightSensor value is <
   {userPreference in the Digital Personality}

   AND a PresenceSensor detects somebody

   AND the LightSensor, the PresenceSensor
    and the LightActuator are in the same room

THEN Turn the LightActuator on
```

This rule is written here in pseudo-natural language for the sake of simplicity; in practice it is expressed in the formal syntax specific to the Jena reasoning engine as shown on Figure 6.
The reasoning engine reads the current light level, the current presence status and the user preferences in the ontology. The premises of the rule are true, so the conclusion must be executed. To determine which rules to apply, Jena uses a classical *forward chaining* reasoning algorithm.
3) Therefore a new statement is added in the ontology: {LightActuator, hasCommand, "on"} (cf. III-C and III-D). In consequence, the framework sends an event to the driver to indicate that the LightActuator must be turned on.
4) The driver commands the light actuator through the Zigbee network. This actually turns the light on.

```
[CMD_LIGHT_ON:
    (?MS RDF:type AMI:PresenceSensor),
    (?LS RDF:type AMI:LightSensor),
    (?LA RDF:type AMI:LightActuator),
    (?R RDF:type ?RT),
    (?RT RDFS:subClassOf AMI:Room),
    (?MS AMI:isIn ?R),
    (?LS AMI:isIn ?R),
    (?LA AMI:isIn ?R),
    (?DP RDF:type AMI:AAL_DP),
    (?DP AMI:isCurentDP ?curDP),
    equal(?curDP,'true'),
    (?DP AMI:low_AAL_LightThreshold ?LLT),
    (?MS CORE:realStateStringValue 'personInside'),
    (?LS AMI:realIntValue ?LMV),
    lessThan(?LMV,?LLT),
    (?F RDF:type AMI:OnOffFunctionality),
    (?LA AMI:hasFunctionality ?F),
    (?C RDF:type AMI:OnCommand)
                         -> (?F AMI:hasCommand ?C) ]
```

Figure 6. Example of rule ("turn the light on") expressed in Jena's syntax.

### B. Diagnosis

At this point, the framework performs diagnosis so as to determine if the action has been executed correctly. The LightActuator is a "light effect" producer; the LightSensor measures the "light intensity" value of "light effect". They are in the same room, so a link between them is deduced automatically. Moreover, if some position reporting system is available, then the physical law associated with "light effect" (that calculates the light intensity) takes into account the actuator-sensor distance. The steps go on like this:

5) The diagnosis framework calculates the *expected light*

*level* at the light sensor by applying the formula associated with "light effect". The result is 120 lux.

6) The light level actually measured by the light sensor is still 80 lux, so the system deduces that there is a failure. The source of the failure (sensor or actuator) is *a priori* known with a limited probability only, but a second sensor in the room may increase it.
7) The system finds it most probable that the bulb is burnt out. An error notification is generated so that the user 1) confirms the cause the problem, and 2) possibly to fixes it (often, even an elderly person is capable of replacing a light bulb). For a discussion on the acceptability of notifications in a home environment, see for instance [13].

### C. Implementation and Results

This experiment uses the standard CBDP framework, with a bundle containing its specific rules. The experiment was conducted in two ways:

- using a simulator of the sensors, actuators and physical environment,
- using physical devices in an actual room.

Figure 7 shows the interface of the simulation environment. The experimenters can act on the light level of the sun, on the motion sensor, and they can also introduce a defect in the light bulb. Both in simulation and in real conditions the system displays a message with the current diagnosis (Figure 8). The tests performed showed that the example runs as expected.

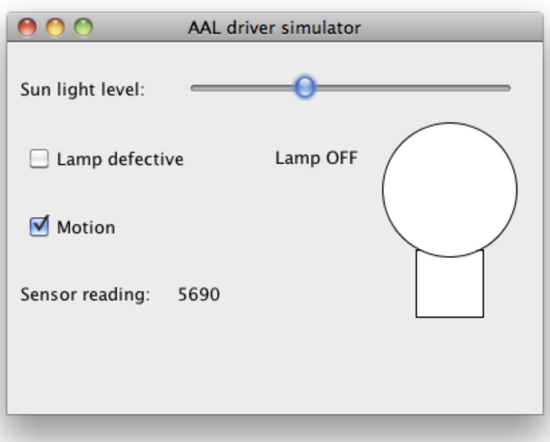



Figure 7. Interface of the simulated environment.

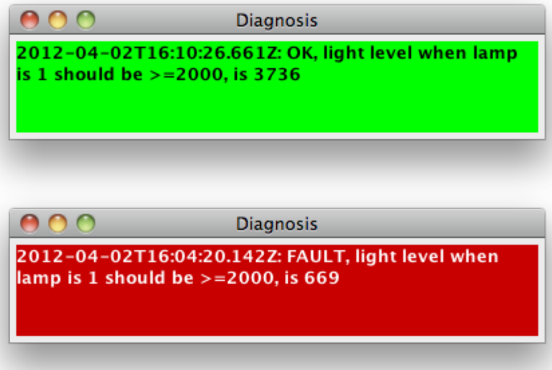



Figure 8. Window showing the results of the diagnosis.

## VI. Related work

Ontologies are often at the heart of ambient-intelligent systems, and especially AAL systems, such as in OASIS [4]. In 2003, CoBrA (Context Broker Architecture) was an ontology-based framework for ambient settings [14]. In 2004, SOUPA (Standard Ontology for Ubiquitous and Pervasive Applications) [15] was one of the first attempts to define an application-agnostic ontology for ambient systems, but it is specifically aimed at agent-based architectures. More recently, Paganelli et al. [16] introduces a tele-health platform, which is based on an ontology for describing context and medical conditions. The SOPRANO project [17], [18] defines a specific ontology that serves as a unifying vocabulary between software components. In our work, the ontology is specifically used to personalize the system: it stores preferences, and contains application-specific modules. Moreover, we not only reason to infer new facts about context as done in many platforms [19], but also to trigger application-specific behavior, and to actually *trigger actions*, i.e., send commands to actuators. This makes the framework flexible and allows the easy integration of additional services such as the diagnosis framework described in Section IV.

Our choice of using the OSGi middleware was motivated by previous successful attempts in the field of ambient intelligence, such as in the AMIGO IST project [20]. CBDP's generalized reliance on ontologies makes the use of OSGi very consistent with the rest of the framework.

Some works focus on ontologies for specific domains. For instance, Hois [21] describes a well-grounded framework for the description of spatial relationships and spatial reasoning. This kind of contributions could be integrated into the CBDP framework, due to the reusable nature of ontologies.

## VII. Conclusion and Future work

We have presented a complete framework that supports the creation of AAL applications. This framework is based on the use of an ontology at the core of the system. This ontology contains application-specific knowledge and stores user preferences ("Digital Personality"). Besides it handles all the run-time information flows: it aggregates sensor data, allows rules to be applied on this data so as to generate commands, stores the commands, and provides the commands to the actuators.

 ISBN: 978-1-61208-235-6 

Using an ontology allows one to specify the behavior of an AAL application in terms of easy-to-write logic rules. These rules can rely on any piece of knowledge present in the ontology, therefore they are not limited in any way by the core ontology that comes with the CBDP framework. Such extensibility is made easy by the use of widespread knowledge engineering standards, namely RDF/OWL.

The other significant contribution of this paper is the diagnosis framework that monitors the run-time behavior of an AAL system by observing changes in the ontology. Currently we take into account only the current state of the system. In reality, the relevant measure might not be the current absolute value of a physical parameter, but rather its *relative evolution*. For instance, when light is switched on, it may be most relevant to consider the *relative increase of the light level*, as the absolute value may vary other time without any action being taken (depending of the intensity of the sun for instance). This prompts us to introduce *dynamics* in the diagnosis framework. Likewise, some physical laws may depend upon quantitative *time* (for instance, the effect of a radiator in an initially chilly room is a slow increase of temperature over time). This is currently being investigated.

We also plan to test such a system in real scale, for example at the homes of elderly people. This will allow us to refine the rules that define the system behavior.

## ACKNOWLEDGEMENT

This work has been performed within the CBDP project, a project co-funded by the European Union. Europe is involved in Région Île-de-France with the European Regional Development Fund.

## REFERENCES

[1] G. Van den Broek, F. Cavallo, L. Odetti, and C. Wehrmann, "Ambient Assisted Living Roadmap," AALIANCE, Tech. Rep., 2009.

[2] CBDP project description. Accessed on 18 July 2012. [Online]. Available: http://projects.celtic-initiative.org/cbdp/

[3] N. Noy and D. McGuinness, "Ontology development 101: A guide to creating your first ontology," Stanford University, Tech. Rep., 2001.

[4] J. Bateman, A. Castro, I. Normann, O. Pera, L. Garcia, and J.-M. Villaveces, "OASIS common hyper-ontological framework (COF)," OASIS Project, Tech. Rep., 2009.

[5] G. Antoniou and F. van Harmelen, "Web ontology language: OWL," in *Handbook on ontologies*, S. Staab and R. Studer, Eds. Springer, 2009, pp. 91–110.

[6] D. Bonino and F. Corno, "DogOnt–ontology modeling for intelligent domotic environments," *The Semantic Web-ISWC 2008*, pp. 790–803, 2008.

[7] Jerry R. Hobbs and Feng Pan. (2006) Time Ontology in OWL. Accessed on 18 July 2012. [Online]. Available: http://www.w3.org/TR/owl-time/

[8] OSGi Alliance, *OSGi service platform, release 3*. IOS Press, Inc., 2003.

[9] J. Carroll, I. Dickinson, C. Dollin, D. Reynolds, A. Seaborne, and K. Wilkinson, "Jena: implementing the semantic web recommendations," in *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*. ACM, 2004, pp. 74–83.

[10] A. Chandra and D. Harel, "Horn clause queries and generalizations," *The Journal of Logic Programming*, vol. 2, no. 1, pp. 1–15, 1985.

[11] Atego. Aonix PERC product description. Accessed on 18 July 2012. [Online]. Available: http://www.atego.com/products/aonix-perc/

[12] A. Mohamed, C. Jacquet, and Y. Bellik, "Real-time Diagnosis of Ambient Environments Using a Modeling of Physical Effects Combined with Temporal Logic," in *Proceedings of UBICOMM 2011*. IARIA, Nov. 2011, pp. 145–151.

[13] M. Vastenburg, D. Keyson, and H. Ridder, "Considerate home notification systems: a field study of acceptability of notifications in the home," *Personal and Ubiquitous Computing*, vol. 12, no. 8, pp. 555–566, 2008.

[14] H. Chen, T. Finin, and A. Joshi, "An intelligent broker for context-aware systems," in *Adjunct Proceedings of Ubicomp*, vol. 3, 2003, pp. 183–184.

[15] H. Chen, F. Perich, T. Finin, and A. Joshi, "SOUPA: standard ontology for ubiquitous and pervasive applications," in *Mobile and Ubiquitous Systems: Networking and Services (MOBIQUITOUS), 2004*. IEEE, 2004, pp. 258–267.

[16] F. Paganelli and D. Giuli, "An Ontology-Based Context Model for Home Health Monitoring and Alerting in Chronic Patient Care Networks," in *Proceedings AINA 2007*. IEEE Computer Society, 2007, pp. 838–845.

[17] M. Klein, A. Schmidt, and R. Lauer, "Ontology-centred design of an ambient middleware for assisted living: The case of SOPRANO," in *30th Annual German Conference on Artificial Intelligence (KI 2007), Osnabrück*, 2007.

[18] P. Wolf, A. Schmidt, and M. Klein, "SOPRANO-An extensible, open AAL platform for elderly people based on semantical contracts," in *AITAmI, ECAI'08*, 2008.

[19] A. Bikakis, T. Patkos, G. Antoniou, and D. Plexousakis, "A survey of semantics-based approaches for context reasoning in ambient intelligence," *Constructing Ambient Intelligence*, pp. 14–23, 2008.

[20] F. Ramparany, R. Poortinga, M. Stikic, J. Schmalenstroer, and T. Prante, "An open context information management infrastructure the ist-amigo project," in *Proceedings of IE 2007, the 3rd IET International Conference on Intelligent Environments*. The IET, 2007, pp. 398–403.

[21] J. Hois, "Modularizing spatial ontologies for assisted living systems," in *Proceedings of the 4th international conference on Knowledge science, engineering and management*, ser. KSEM'10. Springer-Verlag, 2010, pp. 424–435.

# Context-Based Generation of Multimodal Feedbacks for Natural Interaction in Smart Environments

Didier Perroud, Leonardo Angelini, Omar Abou Khaled, Elena Mugellini
Department of Information and Communication Technology
University of Applied Sciences of Western Switzerland
Fribourg, Switzerland
{didier.perroud, leonardo.angelini, omar.aboukhaled, elena.mugellini}@hefr.ch

*Abstract*— **This paper presents a complete system for the generation of multimodal feedbacks in smart environments. It is based on an underlying framework, NAIF, which allows taking into account context information in order to choose the better outputs for the user, improving the effectiveness of feedbacks and the naturalness of the interaction. The extension of NAIF is composed by three main components: a fission engine, a context engine and a context service. Several examples are provided to show the working principles of the developed system. The extended NAIF framework allows setting up applications for the delivery of consistent and quality multimodal feedbacks in heterogeneous environments.**

***Keywords-Multimodal feedbacks; context-aware; smart environment; reasoning engine; multimodal fission.***

## I. INTRODUCTION

Technology is spreading in our homes, many objects become smarter and many devices, such TVs or media-players, have now Internet access. However, most of them offer functionalities that have a limited scope; they often neither communicate with other near devices, nor try to retrieve information from the environment to modify their behavior. Nowadays, many devices can communicate in our homes through protocols like DLNA (Digital Living Network Alliance) [14]; however, the functionality of these protocols is limited to the share of media content. Frequently, the heterogeneity of household devices forces the user to adapt to each system, often without the possibility to choose the best way to communicate with them, for example using different modalities.

Since 1991, Weiser stated the importance of having technologies that disappear to the human sight, integrated in our lives as imperceptible companions [11]. Computers are now smaller, easy to bring with us or embedded in small objects or either in textiles; thus, actually, they are physically disappearing. However, these devices often neither take into account the user perspective, nor they exploit the information from the environment to ameliorate the user perception. In order to weave all these devices together and to create a smart environment, it is necessary to allow them to communicate, to obtain context information and to share embedded resources. These resources can be computational capabilities, information from sensors, stored data, audio and video outputs and actuators.

In this paper, we present a system for the generation of multimodal feedbacks, which takes advantage of the heterogeneous set of devices that can be found in a smart environment. These devices, thanks to their embedded sensors, can contribute also to the generation of context information, which is used to provide consistent and efficient feedbacks to the user. The proposed system is based on the Network Ambient Intelligent Framework (NAIF), a framework for smart environments that has been previously presented in [5]. We have extended NAIF introducing an entity-based context model, a reasoning engine based on first order logical trees and the management of feedbacks using context information. Furthermore the system is able to combine multiple outputs according to the model proposed by Vernier and Nigay [8]. This framework is intended to be easily extensible with new information about context and new heterogeneous devices that can offer input or output resources. Several applications can use these resources at the same time to deliver messages to the user taking advantage of different modalities.

The paper is structured as follows: Section 2 reviews existing projects that are most related to our work. The third section presents the architecture and the implementation details of the system. Finally, we provide three practical examples based on the proposed system, which illustrate its potentialities in different scenarios. The three examples aim to deliver a multi-content message from the well-known social network Twitter, using context information to assure that the multimodal feedback has the best suitability for the conditions of the environment. Future improvements and conclusion end the paper.

## II. RELATED WORK

According to the paper of Dumas et al. [1], the use of natural means of communication facilitates the human interaction with the machine. Communicating using different interaction modalities can be definitely considered as an innate characteristic of human beings. This is generally addressed in the machine through the fusion of inputs modalities and the fission of outputs. Foster defined the multimodal fission as “the process of realizing an abstract message through output on some combination of the available channels” [3]. In our work, we perform multimodal fission in an intelligent environment, which claims to be non-intrusive for the human. Multimodal fission is a research topic that is not often addressed in the scientific community. Although multimodality is a popular topic of research, papers on multimodality address most frequently the fusion process that aims to interpret signals as input. Most systems that use multimodality consist in a natural dialog application, like the European project SmartKom [9], which aims to design a virtual assistant that communicates with the user through gestures and voice. In the SmartKom project, multimodality is limited to a few specific modalities and the fission process cannot be applied to a distributed smart environment. Foster presented in 2003 a survey of multimodal fission [3]. This document is a deliverable of the project COMIC, a dialogue system that uses multimodality. The output generation of the COMIC project is described in [2]. Although the state of the art proposed by Foster is several years old, some key points of the fission process are still very relevant. Foster explains that the fission process can be separated into three parts. The first part is the content selection and the structuring of the information. The second part is the selection of modalities based on context knowledge. The final part deals with the coordination of the outputs in order to provide consistent feedbacks. Our system is not focused on the content selection, but implements a selection of modalities based on the context retrieved from an intelligent environment.

Vernier and Nigay presented in [8] a conceptual space about the output combination of modalities. The authors explain that multimodal outputs can be generated according to five different compositions. These five compositions can be further interpreted according to different aspects: time, space, articulation, syntax and semantics. Our work is based on this conceptual space to perform the task of output coordination in the fission process. The fission process is performed as part of an intelligent environment; we propose a composition of output modalities in time and space and the control of generators according to the semantic aspect of composition. The possible compositions of these aspects are presented in Section 3.

The work presented in this article is an extension of the previous approach that we proposed in [5]. This approach was intended to propose a solution for the generation of context-aware multimodal feedbacks based on NAIF, which aims to set up and operate a smart environment. The main components introduced in that article were a context service, a reasoning engine for context and a fission engine. Although a simple working prototype has been already presented, many parts of this concept were neither implemented, nor detailed. Indeed, the fission process and the reasoning techniques were not defined and no context modeling was proposed. Our contribution aims to overcome these different lacks. We propose a context model and a reasoning engine that aim to exploit the context of a smart environment. We also present a multimodal fission process based on our context management solution.

The multimodal fission process proposed in this paper is based on the context of the environment. One of the main challenges of exploiting context information is to model it. The scientific community is very active on this topic. In the literature we find generally three mains trends about context modeling. The first trend is ontology-based. Wang et al. recommend the use of this technique in [10]. Ontologies also offer the ability to use inference to reason on context. However, Henricksen et al. use an object-oriented model in their work [4]. This second technique facilitates the implementation of context management, thanks to its similarity with objects-oriented programming languages. Finally, Ranganathan et al. proposed a modeling technique based on predicates [6]. This technique allows easily applying advanced mathematical reasoning on context information. Each mentioned modeling technique has its strengths and weaknesses. As shown by Ranganathan et al. in [7], an interesting approach is to use several techniques simultaneously. Indeed, they used an ontology model to share vocabulary and a predicate model for the implementation. A Survey of context modeling and reasoning techniques can be found in [16]. In our work, we have chosen to use an object-oriented model to facilitate the implementation of context management. In fact, the context management solution is integrated into a Framework that must be used by several developers that have their own representation of the context.

## III. SYSTEM ARCHITECTURE

NAIF operates in three layers: the inter-communication

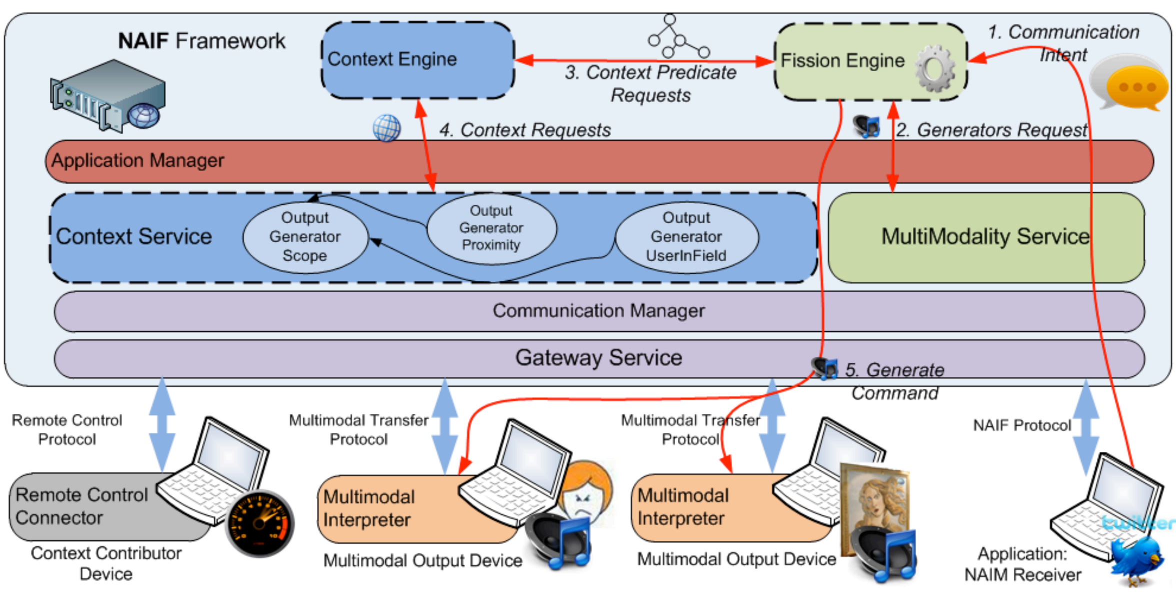



Figure 1. NAIF Extended framework for context-aware multimodal feedback; dashed lines show improved components.

layer, which includes the gateway and the communication manager, an interworking layer, which offers services and shared resources to the applications, and the application layer, which can handle applications running on the devices connected to NAIF. Further details about the architecture could be found in [5]. In order to generate multimodal feedbacks suitable for the context of the environment, NAIF has been improved in several parts, depicted in Figure 1. First, the Context Service for context generation has been ameliorated with an object-oriented entity model for context representation. Second, the Context Engine has been extended to handle requests for context retrieval, using first order logical trees. Finally the Fission Engine has been improved, taking advantage of the new Context Engine, and introducing spatial, temporal and semantic compositions of the feedbacks. NAIF is extensible to several applications and heterogeneous devices that want to communicate information to the user or offer proper data for the generation of the context. These devices can benefit from outputs available in the environment, publish sensor information and output capabilities according to the three respective protocols depicted in Figure 1: the NAIF Protocol with communication intents, the Remote Control Protocol and the Multimodal Control Protocol. Both the Fission Engine and the Context Engine run in the application layer, thus they can be easily modified or replaced without affecting the operation of the whole framework. In the following paragraphs (A, B and C), the principles of operation of the main components for context-based generation of multimodal feedbacks are explained. The five steps necessaries to the generation are shown in Figure 1.

An application that wants to produce an output has to send a *communication intent*, a structured message that specifies to the Fission Engine the type of media, an eventual recipient and the composition of information, if more media are present (Step 1). Multiple media are handled according to three of the five composition aspects presented by Vernier and Nigay [8]. An adaptation of their model for our system is shown in Figure 2. Communication intents can specify temporal composition parameters, i.e., the delay between each message, and a spatial composition, for example if two media should be played in adjacent or opposite displays. Finally, the semantic attribute specifies if the media that are present in a communication intent are semantically complementary, thus if both must be reproduced in order to correctly deliver the information. When a communication intent is received, the Fission Engine retrieves the list of the available Output Generators from the MultiModality Service (Step 2). Generators are filtered according to their capabilities of displaying media, to the scope or recipient of the message, spatial composition and context requirements. Each time that context information is needed, a logical tree request is formulated by the Fission Engine and sent to the Context Engine (Step 3), which retrieves data from the Context Service (Step 4). Finally, generation commands are sent to the chosen outputs, which, if necessary, will be locked to prevent further uses by other applications while displaying the feedback (Step 5). Temporal composition attributes are used to synchronize outputs, if specified. Further details are presented in the following paragraphs.

## A. *Context Model and Generation*

The context is represented by an object-oriented model and is maintained up-to-date by the Context Service. Each Entity is defined in a stand-alone XML Schema that contributes to the global model. The choice of an object-oriented approach instead of an ontology model is made to ensure rapid development and easy extension even for novice programmers. The model is based on Context Entities and a Context Entity Manager for each entity type. The Context Entity Manager generates and updates the corresponding entities following three different strategies. These strategies are implemented according to the context acquisition methods proposed by Mostefaoui et al. in [17]. For Sensed Context Entities, e.g., Luminosity or Noise, it retrieves necessary data from the NAIF Remote Control Service, which collects measured data from sensors embedded in the environment or in the connected devices. For Knowledge Based Context Entities the Context Entity Manager obtains data from local (e.g., sensors positions) or online (e.g., weather) databases. The last type of Context Entity Manager handles Derived Context Entities combining

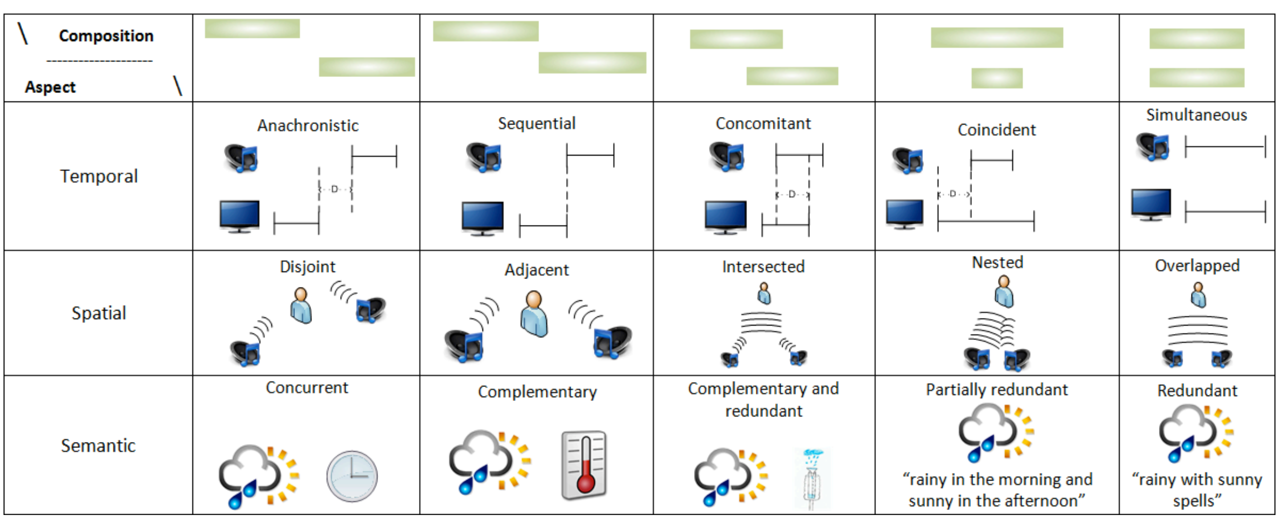



Figure 2. Spatial, temporal and semantic composition of outputs according to the model of Vernier and Nigay.

and elaborating information from one or more entities. Adding a new Context Entity to the global model is as easy as defining it in an XML Schema and implementing the Context Entity Manager, which retrieves data from sensors and transform them to high level thresholds. These thresholds, e.g., high, low or medium, are easier to understand for an end-user and they are fixed by the programmer during the system configuration. For the purpose of the adaptive multimodal feedback generation, we have defined a specific context model that describes the state of the Output Generators. This basic model is showed in Figure 1 in the Context Service module. The first type of entity, Output Generator Scope, describes a generator and specifies the type, the range or scope, and the position. These entities belong to the Knowledge Based Context Entity type and will not be automatically created by the NAIF Framework, but generally must be defined by the programmer. The Output Generator Proximity Entity aggregates information of two Output Generators and determines if they are reciprocally near or far. Similarly, the Output Generator UserInField, determines if a person is within the scope of an Output Generator or not. Only available Output Generators are present in the context state. The Person Entity used for the tests contains just the name and the ID of the associated RFID tag. The model obviously involves also other entities like Luminosity, Noise, etc., which will not be described in detail.

### B. *Context Engine*

The Context Engine allows handling complex queries on the context state, using a protocol formalized by first order logical trees. Both requests and results are formalized by a tree of entities from the context state, shown in Figure 3. A first order logical tree is a logical expression (AND/OR), which involves entities values, or recursively, other sub-trees. Each leaf of the tree is represented by an entity state. In the request, desired entities states are specified; to elaborate the query result, the correspondent entities are retrieved with the desired states from the Context Service: if there are no corresponding entities with the specified state, the leaf has a false logical value; otherwise the corresponding entities are put in the result tree and the logical value is true. Providing a tree of entities as a result is very important to retrieve information about selected entities. Moreover, each branch of a tree can be marked as *required* or as *optional*. If a branch is optional, it will not affect the truthfulness of the parent expression, however the corresponding entities, if their expressions are true, will be included in the query result. Optional branch of predicates are also useful to

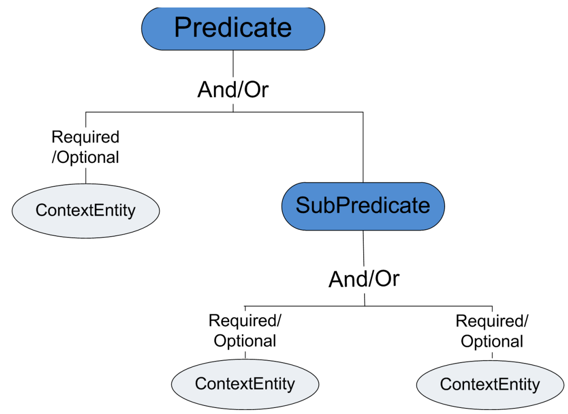



Figure 3. First-order logical trees for context requests.

retrieve information that can be used to ameliorate the quality of feedbacks when basic constraints are verified. An example of context request is presented in Figure 4. An application wants to send a video message to Bob and Alice. The Fission Engine build the request asking desired context states for four different entities, but only two of them are marked as required. The example supposes that in the Context Service the Noise state is low, the Luminosity is high and that Bob and Alice are in proximity of the TV. It is worth noting that the Luminosity entity state is not included in the response because it is marked as optional and has a false logical value. The Context Engine can accept request also from other applications that want to directly access context information without delegating the work to the Fission Engine.

### C. *Fission Engine for Multiple Feedbacks*

The role of the Fission Engine is to choose a proper output and modality for high-level media, e.g., text that can be displayed to the user either as visual feedback or as aural feedback, using Text To Speech (TTS). The Fission Engine makes this choice according to the context information and an eventual spatial composition. The overall selection process has been already described in Figure 1. In this paragraph, the selection process of the Output Generators and their management during the generation of feedbacks is

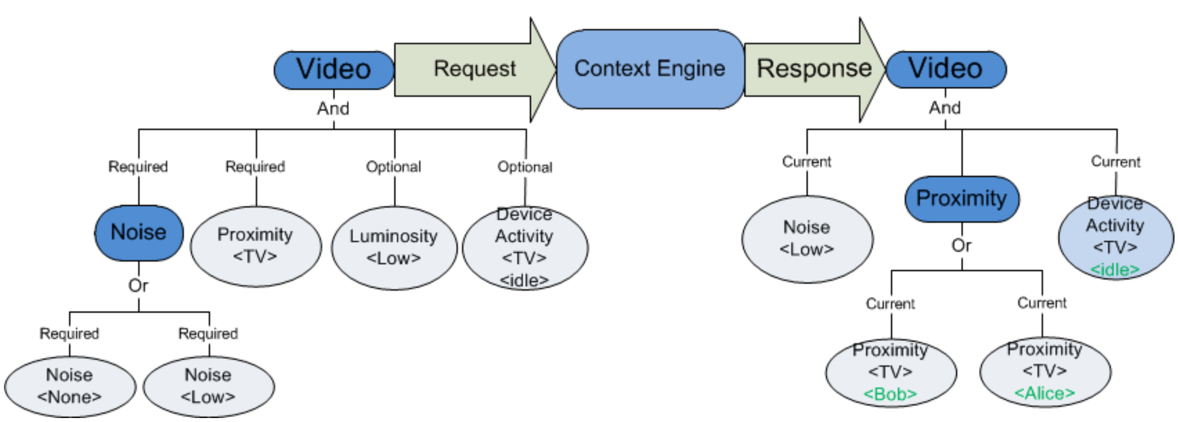



Figure 4. Example of context request.

detailed.

Figure 5 resumes the workflow of the Fission Engine and shows the related exit points, which occur when one or more requirements are not satisfied and the generation of the outputs is not possible. In few cases, detailed in this paragraph, the generation is not stopped even if some constraints or requirements are not fulfilled. First of all, the Fission Engine checks if the types of media required by the communication intent are in the list of compatible types (for example Text, Images, Emotions, etc.). Then, the available generators are retrieved from the Multimodal Service and filtered according to the media types of the communication intent. Output Generators are further filtered according to their scope: for messages without a recipient, global generators (available for all people in the room) are chosen. Otherwise, if a recipient is specified, the Fission Engine formulates a request to the Context Engine in order to determine which Output Generators are in the same scope of the recipient. The current implementation can handle only a single recipient. Applications that want to send messages to multiple users should replicate the communication intent for each recipient. For global scope multiple feedbacks, spatial combination is taken into account and the available spatial combinations are then filtered according to the context requirements. The spatial combination for multiple feedbacks is chosen among all possible permutations of Output Generators. This technique is acceptable for small environments but becomes very inefficient when many Output Generators are present and should be improved for this latter case. The choice of the better spatial combination (or of the single Output Generator for communications composed by only one media) is made according to context requirements for each type of modality for that media. Requirements are generally defined by the programmer in a database and are used to construct the context requests as defined in the previous paragraph. When there are no generators with an adequate context, the behavior of the Fission Engine changes according whether the communication intent has more than one media or only one.

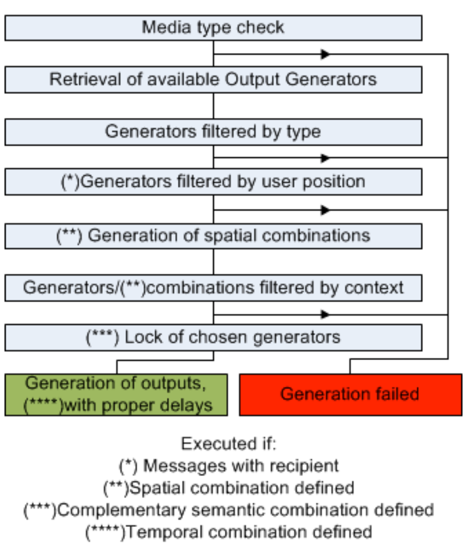



Figure 5. Fission engine workflow.

In the first case, the Fission Engine sends the output commands to the chosen generators even if the context does not respect the defined requirements, because the multiple nature of the communication can probably help the user to understand the message even if some outputs are perturbed (Assumption 1). If only one media is present, instead, we designed the Fission Engine to make the communication fail when none of the generators has an adequate context. Thus, we believe that using only one perturbed generator it is not possible to correctly deliver the information to the user (Assumption 2). The Assumption 1 is probably correct only if the semantic composition of the message is *redundant*. In the case of *complementary* information, the Fission Engine should drop the communication intent. However, in the current implementation of the system, the Fission Engine is not able to choose generators according to *semantic combination* of the information, nor to assign more generators to the same piece of information, in order to augment the quality of the feedback in the case that all the available generators are perturbed. However, semantic information is taken into account during the generation of feedbacks. In fact, Output Generators that display non-redundant information are locked by the Fission Engine in order to avoid that other applications could interrupt the output generation. Delays are applied to the generation of multiple messages in order to respect the desired temporal combinations.

## IV. Scenario

The main objective of the proposed system is allowing a seamless interaction in our homes using existing technologies, just interconnecting current available devices through the NAIF Framework. We developed an application, Natural Ambient Internet Messaging (NAIM), which uses the Context Engine and the Fission Engine of NAIF to deliver proper feedbacks to the users in a smart environment. NAIM is a simple connector to the popular social network Twitter. For our purpose, i.e., testing the multimodal fission of NAIF, we developed only a receiver that is able to interpret messages written on a dedicated account with a proper syntax. The syntax allows defining in a tweet several messages at one time, which can involve different type of media: text, emotions and images. Moreover, the composition of these messages can be specified according to the spatial, temporal and semantic axes. This is done by adding special text commands at the end of the tweet as shown in Figure 6.


Figure 6. A local scope message (green, in red the recipient) with emoticon (orange) and a global message (green) with spatial and temporal composition (yellow).

In the following paragraphs, we analyze three different scenarios based on the NAIM Receiver that show how the

 ISBN: 978-1-61208-235-6 


Figure 7. Scenario: 1) TV; 2) Beamer; 3) Aphrodite; 4) Laptop.

multimodal context-aware generation can benefit to the Human Computer Interaction within a smart environment. The three scenarios have been implemented and tested in a smart living room where several devices were connected through NAIF. Each device was connected through an Ethernet or Wi-Fi connection to another PC that was running the NAIF base Framework (Gateway), the Context Engine, the Fission Engine and the NAIM Receiver. Obviously, these three latter components are applications that could run on any computer connected to the NAIF Gateway. In particular, as shown in Figure 7, we used a media center connected to a TV, a Personal Computer connected to a beamer and a laptop for local scope messages. Context information is obtained using popular Phidgets [15], in particular a microphone and a luminosity sensor. An RFID reader has been used to detect the presence of a target user.

Once a tweet is received by the NAIM Receiver, the time necessary to generate feedbacks was within the order of magnitude of the transmission time of the messages over the local network. The major delay between an input and output of this demonstrator is related to the time necessary to Twitter servers to store and send back the tweet to the NAIM Receiver through their services.

The NAIM Receiver decomposes the tweet and builds a communication intent composed of more than one media. Three different media have been implemented in this scenario: text, images and emotions. For text messages, two output modalities are available: a visual one, which displays the text in a window, and an aural one, which uses Text To Speech (TTS). Obviously, images can be displayed only with a visual feedback. Emotions, instead, are represented either with visual smiley or by animating a physical painting called Aphrodite [12]. Aphrodite is able to represent emotions through sounds and expressions of the face in the painting. The first media of a communication intent is always the text "New message received by SenderName", where SenderName is the Twitter account that published the message.

### A. *Global Text Message in a Noisy Environment*

The first scenario involves a global message to be displayed to all the users, e.g., an alert message. This message, which is sent as text media, will be displayed as an aural feedback using TTS if the noise of the environment is low or as a visual feedback on available screens otherwise. When the tweet is received by the NAIM Sender, it generates a communication intent that includes two text media: the standard text with the sender of the tweet and the content of the tweet. Then, the Fission Engine requests global scope Output Generators that are able to display text, i.e., in our scenario, the TV and the beamer. For each generator, two modalities are available to display text: an audio modality (TTS) and a visual modality (text displayed in a window). In this scenario no constraint has been set for the luminosity; the TV audio generation (TTS), instead, had a constraint on the noise measured by the associated microphone. This constraint has not been set for the beamer, which was paired with more powerful speakers. Further, the Fission Engine controls if the constraint is verified and chooses the best output generator for each of the two media. When only the TV is turned on and there is noise in the room, both messages will be displayed with visual feedbacks; otherwise, with no noise, the system generally prefers to use both visual and aural feedbacks, assigning one of this modality to each of the two messages. If also the beamer is turned on, the Fission Engine will choose the TV for the text display and the beamer for TTS. Introducing a new constraint on the context of the room in order to provide visual feedbacks on the beamer only if the luminosity is low will lead the Fission Engine to use TTS on the beamer or to display both messages as text on the television. This example shows how the system can automatically improve the quality and the efficacy of the feedbacks.

### B. *Messages Addressed to a Mobile User*

The same scenario is now considered with messages with a specified recipient. The user is authenticated using an RFID tag that is recognized by a computer with a local scope Output Generator. The target application would like to provide feedbacks to the user even when he or she is moving in the home, choosing each time a local Output Generator and the best output modality according to noise or luminosity information. The Fission Engine in this case analyzes the recipient field of the communication intent and finds proper local generators according to the information of the Output Generator UserInField entity.

### C. *Multimedia Message with Context-adaption*

Previous examples involve media of only one type: text. In order to demonstrate the multimodal fission of a message, in this scenario emotions (emoticons) have been included. While text will be displayed either with visual or aural feedbacks, emotions will take advantage also of the aforementioned Aphrodite system, which is able to display them changing facial expressions in the painting. However, this picture cannot be seen with low luminosity: in this latter case the system will prefer to display the emotion as a smiley on a screen, for example the beamer. On the other hand, a smiley could not be displayed on the beamer if the luminosity of the room is high. Similarly to the first scenario, it is possible to define these context requirements in order to allow the Fission Engine to choose the best Output Generators.

## V. CONCLUSION AND FUTURE IMPROVEMENTS

This paper presented a system that introduces an object-oriented model to represent context, a Context Service to maintain up-to-date it, a Context Engine to handle requests on the context state and verify desired conditions and a Fission Engine to select the best outputs for a given communication intent. An application that delivers messages from a Twitter account to the user has been developed in order to demonstrate the working principle of the system. NAIF with the management of multimodal feedbacks according to context information allows the delivery of high quality feedbacks within a smart environment, taking advantage of input and output resources available on present devices. The system allows to several applications to share these resources dynamically, building up a model of the context and choosing proper resources according to context information.

Obviously some limitations are present: NAIF lacks a system of coordinates that could allow an accurate definition of the position of the Output Generators, of the user and of the context sensors. By now, the positions of the generators are defined by giving a name to the zones where they are located, with no definition of proximity and relative distance. A system of coordinates and the tracking of users as that shown in [13] should be included in order to enhance the capabilities of NAIF and the quality of the feedbacks with spatial composition. Moreover, types of media, context requirements and Output Generators capabilities are defined with XML files, which are not very intuitive to be configured by end-users. A graphical user interface should be implemented in order to allow the user to easily set up a new system based on NAIF.

## REFERENCES


[1] B. Dumas, D. Lalanne, and S. Oviatt, “Multimodal interfaces: a survey of principles, models and frameworks”, Human Machine Interaction, 2009, pp. 3-26,

[2] M. E. Foster, “Producing multimodal output in the COMIC dialogue system”, Proc. of the W3C Workshop on Multimodal Interaction, 2004

[3] M. E. Foster, “State of the art review: Multimodal fission”, COMIC project Deliverable 6.1, 2002 [retrieved: July, 2012]

[4] K. Henricksen, J. Indulska, and A. Rokotonirainy, “Modeling context information in pervasive computing systems”, Proc. of Pervasive, 2002, pp. 79-114.

[5] D. Perroud, L. Angelini, E. Mugellini, and O. A. Khaled “Context-aware multimodal feedback in a smart environment”, AMBIENT 2011, pp. 1-6.B.

[6] A. Ranganathan, J. Al-Muhtadiand, R. H. Campbell, “Reasoning about uncertain contexts in pervasive computing environments”, IEEE Pervasive Computing, 2004, pp. 62-70.

[7] A. Ranganathan, R. E. McGrath, R.H. Campbell, and M. D. Mickunas, “Ontologies in a pervasive computing environment”, Proc. of the Workshop on Ontologies and Distributed Systems, 2003,

[8] F. Vernier and L. Nigay, “Multimodal interfaces: composition and characterizazion of output modalities” (Interfaces multimodales: composition et caractérisation des modalités de sortie), Proc. of ErgoIHM'2000, 2000, pp. 203-210.

[9] W. Wahlster, “SmartKom: fusion and fission of speech, gestures and facial expressions”, Proc. of the 1st Int. Workshop on Man-Machine Symbiotic Systems, 2002, pp.213-225.

[10] X. H. Wang, D. Q. Zhang, T. Gu, and H. K. Pung, “Ontology based context modeling and reasoning using OWL”, Proc. of the 2nd IEEE Pervasive Computing and Communications, 2004, pp. 18-22.

[11] M. Weiser, “The computer for the 21st century”. Scientific American, 1991, 265 (3), pp. 94-104.

[12] Aphrodite, a head capable of reproducing one dozen of human expressions http://www.aphrodite-glance.ch, [retrieved: July, 2012]

[13] Mori, T., Suemasu, Y., and Noguchi, H., “Multiple people tracking by integrating distributed floor pressure sensors and RFID system”, *Systems, Man and Cybernetics*, (2004), pp. 5271-5278.

[14] Digital Living Network Alliance, “DLNA Home Networked Device, Interoperability Guidelines v1.0,” Jun. 2004 [retrieved: July, 2012]

[15] Phidgets, Products for USB Sensing and Control, http://www.phidgets.com/ [retrieved: July, 2012]

[16] C. Bettini, O. Brdiczka, K. Henricksen, J. Indulska, D. Nicklas, A. Ranganathan and D. Riboni, “A survey of context modelling and reasoning techniques.” Pervasive and Mobile Computing 6, 2010, pp.161-180.

[17] G.K. Mostefaoui, J. Pasquier-Rocha, and P. Brezillon, “Context-Aware Computing: A Guide for the Pervasive Computing Community”, Proc. of the IEEE/ACS International Conference on Pervasive Services, 2004. pp 39-48.

# A Multi-Sensor Approach for Activity Recognition in Older Patients

Carlos Fernando Crispim Junior[1], Francois Bremond[1,2]
[1]Institut de Recherche en Informatique et en Automatique
[2]EA CoBTeK Université de Nice Sophia Antipolis
Sophia Antipolis, France
{carlos-fernando.crispim_junior,francois.bremond}@inria.fr

Veronique Joumier[2,3]
[3]Centre Mémoire de Ressources et de Recherche de Nice
Nice, France
veronique.joumier@inria.fr

***Abstract* — Existing surveillance systems for older people activity analysis are focused on video and sensors analysis (*e.g.*, accelerometers, pressure, infrared) applied for frailty assessment, fall detection, and the automatic identification of self-maintenance activities (*e.g.*, dressing, self-feeding) at home. This paper proposes a multi-sensor surveillance system (accelerometers and video-camera) for the automatic detection of instrumental activities of daily living (IADL, *e.g.*, preparing coffee, making a phone call) in a lab-based clinical protocol. IADLs refer to more complex activities than self-maintenance which decline in performance has been highlighted as an indicator of early symptoms of dementia. Ambient video analysis is used to describe older people activity in the scene, and an accelerometer wearable device is used to complement visual information in body posture identification (e.g., standing, sitting). A generic constraint-based ontology language is used to model IADL events using sensors reading and semantic information of the scene (*e.g.*, presence in goal-oriented zones of the environment, temporal relationship of events, estimated postures). The proposed surveillance system is tested with 9 participants (healthy: 4, MCI: 5) in an observation room equipped with home appliances at the Memory Center of Nice Hospital. Experiments are recorded using a 2D video camera (8 fps) and an accelerometer device (MotionPod®). The multi-sensor approach presents an average sensitivity of 93.51% and an average precision of 63.61%, while the vision-based approach has a sensitivity of 77.23%, and a precision of 57.65%. The results show an improvement of the multi-sensor approach over the vision-based at IADL detection. Future work will focus on system use to evaluate the differences between the activity profile of healthy participants and early to mild stage Alzheimer's patients.**

***Keywords – Event Streams Fusion; Instrumental Activities of Daily Living; Dementia; Elderly***

## I. INTRODUCTION

The growth of older people population in past years has highlighted the importance of the development of tools that could extend their independent living, and to improve the diagnosis and treatment of age-related diseases. Information and Communication Technologies (ICT) have proposed wearable sensors [1-2], smart homes, and video surveillance systems [3-7] to address older people surveillance, detect emerging patterns of frailty and falling down events, and to support the early diagnosis of aging related diseases (*e.g.*, Alzheimer, Parkinson, Mild Cognitive Impairment - MCI). People surveillance is generally based in smart home approaches, which analyzes older people daily living activities using a set of heterogeneous sensors placed at home appliances (*e.g.*, infrared presence sensors, contact sensors, temperature sensors, microphones, and video cameras). Their goal is generally to identify falling down events or emerging patterns of frailty by the detection and analysis of a set of self-maintenance activities known as Activities of Daily Living (ADL[8], *e.g.*, hygiene, toilet use, eating, resting, and dressing) [10-12]. Although a multi-sensor approach enriches the quantity of data about the person daily routine, the multiple sources of readings also increase the complexity of data analysis process, as it is necessary to choose the relevant sensor to detect the activity of interest (disregarding the existent data storage issues).

Video-monitoring systems are an alternative to the fully sensor equipped smart-homes, as they can replace or be used in parallel to reduce the number of sensors necessary to describe the overall activity of a person. Existing applications vary from fall detection to ADL detection at constrained environments [3-7]. An ontology-based approach was shown for modeling the context of human status (*e.g.*, body posture) and the environment context (semantic information about the scene). The models use information provided by a set of cameras for person detection, and accelerometer devices attached to objects of daily living for environment events triggering (*e.g.*, TV remote control or doors use). A rule-based reasoning engine is used for processing and combining both models types at activity detection level. The ontology tries to solve the semantic gap among the human activities (and scene context) and the sensors raw signals [28]. A Fuzzy logic scheme [27] was also proposed to cope with multiple sensor fusion at activity analysis in a smart home. Audio, infrared sensors and a wearable device (acquiring physiological signals, like, ECG, and body posture) are combined to infer ADLs events.

Although daily activity surveillance can support the analysis of medium- to long-term patterns of activity, in-lab clinical protocols supported by ICT have been recently proposed to analyze a person performance in specific activities (like ADL), that could highlight emerging symptoms of a certain diseases. For example, wearable devices have been used to assess older people motor functions performance in a gait analysis test. Patients wear a chest or fist sensor which automatically extracts parameters about their gait performance (e.g., stride length and cadence, vital signals). These parameters are used to identify disturbances in gait patterns that could be associated to emerging symptoms of Dementia (*e.g.*, Alzheimer, MCI, and Parkinson) [16-18, 22].

This paper proposes a multi-sensor surveillance system using heterogeneous sensors data (video and accelerometer recordings) to identify activities during a medical clinical protocol. An ontology approach is used to model the semantic information about activities and the clinical protocol tasks. Accelerometer data is combined with visual information to identify the body posture of the person in situations where video camera data is not sufficient due to occlusion or scene characteristics that affect people detection on the scene.

The set of activities analyzed differ from previous studies as Instrumental Activities of Daily living (IADL) are used instead of ADL. IADL set of activities have been recently discussed as a better estimator for the evaluation of emerging neuropsychiatric symptoms, as they are associated with tasks that require independence, organization, judgment and sequencing abilities (*e.g.*, using the phone, shopping groceries, organizing medications, and managing personal finances) [8].

The system prototype is tested in an ongoing medical clinical protocol under development by the Memory Center of Nice Hospital, which intends to quantitatively and objectively assess older people executive functions and functional abilities. The long-term goal of this approach and the clinical protocol is to differentiate early stage Alzheimer's patients from healthy participants.

The main contributions of the present work are the use of accelerometer-based and visual-based information for posture recognition. The sensors data are combined using an ontology language similarly to [28], but the set of sensors used is different. The IADLs recognition instead of ADLs as previous works discussed is stated as a contribution for the clinical domain. Next sections will present the proposed multi-sensor approach, the clinical protocol where the proposed system is tested, and the preliminary results.

## II. MATERIALS AND METHODS

### *A. Clinical Protocol*

Participants aged more than 65 years are recruited by the Memory Center (MC) of the Nice Hospital. Inclusion criteria of the AD group are: diagnosis of AD according to NINCDS-ADRDA criteria and a Mini-Mental State Exam (MMSE) [19, 23] score above 15. AD participants which have significant motor disturbances (per the Unified Parkinson's Disease Rating Scale) are excluded. Controls participants were healthy in the sense of behavioral and cognitive disturbances.

In the designed clinical protocol the participants have to undertake a set of physical activities and IADLs in a Hospital observation room furnished with home appliances. The clinical protocol activities are divided in three scenarios are as follows:

(1) Scenario 01 - Directed activities (10 minutes): it intends to assess kinematic parameters about the participant gait profile (*e.g.*, static and dynamic balance test, walking test);

(2) Scenario 02 - Semi-directed activities (15 minutes): it aims to evaluate the degree of independence of the participant by taking and organizing a list of daily living activities (IADLs) within 15 minutes. Participant is alone in the room with the list of activities to perform, and he/she is advised to leave the room only when he/she has felt that the required tasks are completed;

(3) Undirected ("free") activities (5 minutes): it aims to assess how the participant spontaneously initiates activities and organize his/her time.

The proposed monitoring system is tested with the video recordings of the Semi-directed scenario (Scenario 02). Table 1 shows the set of IADLs that the participant is instructed to perform. Results using a previous version of the clinical and a video-based approach for activity detection be seen in [19].

TABLE 1. CLINICAL PROTOCOL – SCENARIO 02

| |
|---|
| Watch TV, |
| Make tea/coffee, |
| Write the shopping list of the lunch ingredients, |
| Answer the Phone, |
| Read the newspaper/magazine, |
| Water the plant |
| Organize the prescribed drugs inside the drug box according to the daily/weekly intake schedule, |
| Write a check to pay the electricity bill, |
| Call a taxi, |
| Get out of the room. |

### *B. Data recording*

Experimental data is recorded using a 2D video camera (AXIS®, Model P1346, 8 frames per second), a 3D camera (Kinect® sensor), a motion sensor (MotionPod®), and ambient audio microphone (Tonsion, Model TM6, Software Audacity, WAV file format, 16bit PCM/16kHz). The proposed approach is tested with 9 participants (healthy: 4, MCI: 5).

MotionPod® is a proprietary device that provides estimation about a person posture based on the readings coming from accelerometers and gyroscopes sensors. MotionPod® is attached to Patient chest using a chest strap accessory. As Video and MotionPod® recordings are acquired using different computers, these computers had their internal clock reference automatically synchronized using a network service (Network Time Protocol) to avoid time issues.

### *C. Multi-sensor surveillance system*

The proposed multi-sensor system is divided in two main components: the vision and the event detection components. The video record is processed by the Vision Component which is responsible to detect mobile objects and classify them according to a set of objects of interest (*e.g.*, person). Once a person is detected on the video frame under analysis, the event module component analyzes his/her activity in relation to the activity models previously defined by an expert using an ontology language. Posture events coming from MotionPod® data are taken in account only at the event component level. Fig. 1 illustrates a diagram of the system architecture and sensors contribution for each module.

 ISBN: 978-1-61208-235-6 

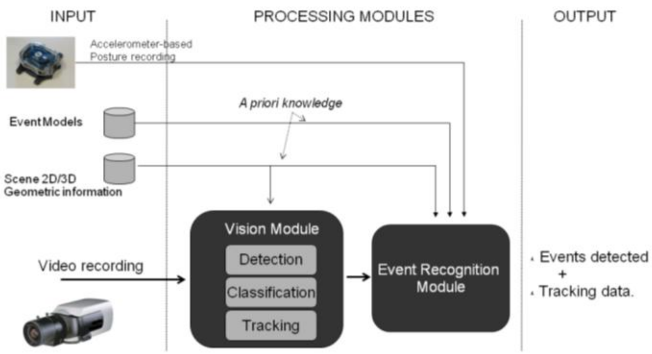



Figure 1. System architecture

*1) Vision Component*

The vision component has been developed using a modular vision platform locally developed that allows the test of different algorithms for each step of the computer vision chain (*e.g.*, video acquisition, image segmentation, physical objects detection, physical objects tracking, actor identification, and actor events detection). The vision component extracts the objects to track from the current frame using an extension of the Gaussian Mixture Model algorithm for background subtraction [26]. People tracking is performed by a multi-feature tracking algorithm presented in [20], using the following features: 2D size, 3D displacement, color histogram, and dominant color. Fig. 2 shows an example of the vision component output. Rectangle envelope highlights the detection of a person in the scene. Blue dots represent previous positions of the detected person in the scene.

Figure 2. Multi-sensor activity recognition system screen. Rectangle envelope highlights the detection of a person in the scene. Blue dots represent previous positions of the detected person in the scene.

*2) Event Detection Component*

Event detection component uses a generic constraint-based ontology language to model and recognize activity events. Event models are described using *a priori* knowledge about the experimental scene (scene model), the clinical protocol activities (event models), attributes of the physical object tracked by the vision component (*e.g.*, spatial, temporal, kinematic and appearance properties), and posture estimations coming from MotionPod® sensor.

The *a priori* knowledge about the experimental scene consists of the decomposition of a 3D projection of the scene floor plan in a set of spatial zones which represent contextual information about the clinical scenario (*e.g.*, zoneTV, zoneArmChair, zoneOfficeDesk). The constraint-based ontology is declarative and intuitive language as it is based in natural terminology, allowing domain experts to easily define and modify the event models.

Event models are hierarchically categorized according to their complexity (primitive or complex) and the information they model (the current state or time-based event). Primitive states refer to a specific value in the property(ies) of a tracked physical object (*e.g.*, the participant). Primitive events refer to a change(s) in the value(s) of an object property(ies) (*e.g.*, change of posture). Composite/Complex events are defined as a set of primitive events and/or states. More details about the generic ontology language used are seen in [25].

Table 2 shows an example of the complex event called "Person standing and using OfficeDesk". The ontology event description is divided among four components: the Physical objects involved in the event model (an object classified as a Person – p1, a contextual zone – z1, and a contextual object – eq. 1); the Components (or sub-events) that must be detected for the Event Model identification (*e.g.*, Person_using_OfficeDesk); the constraints which components and physical objects attributes must hold (e.g., minimal duration of the event "Person using OfficeDesk"); and an Alarm category used as reference to a specific treatment that could be performed in case of the event is detected.

TABLE 2. EVENT MODEL OF PERSON_STADING_AND_USING_OFFICEDESK

```
CompositeEvent(Person_standing_ and_using_OfficeDesk,
   PhysicalObjects(
               (p1 : Person), (z1 : Zone),  (eq1 : Equipment))
   Components(
               (c1 : CompositeEvent Person_using_OfficeDesk(p1, z1, eq1))
               (c2:  PrimitiveState     Person_standing_MP ( p1) )
   )
   Constraints((duration(c1) > 2)
   Alarm ((Level : NOTURGENT))
)
```

Table 3 describes the primitive state called "Person using OfficeDesk". "Person using Office Desk" event is modeled based on the participant presence inside the contextual zone "OfficeDesk", and in the distance between participant 3D bounding box projection on the scene floor and the parallelepiped modeling the real desk.

TABLE 3. EVENT MODEL OF PERSON_USING_OFFICEDESK

```
CompositeEvent(Person_using_OfficeDesk,
     PhysicalObjects( (p1 : Person), (z1 : Zone), (eq1 : Equipment))
     Components(
               (c1: PrimitiveState Person_Inside_Zone_OfficeDesk(p1, z1))
               (c2: PrimitiveState Person_closeTo_OfficeDesk(p1, eq1))
  )
  Alarm ((Level : URGENT))
)
```

*3) Multi-sensor event fusion*

Sensor fusion can be performed at different levels, *e.g.*, data level (raw data combination), feature level (fusion of observation data from different sensors), or at decision level (fusion of events detected by different sensors) [24]. The multi-

 ISBN: 978-1-61208-235-6 

sensor fusion herein presented is performed at decision level. One advantage of sensor fusion at this level is the possibility of changing the sensor type without the need of change of event models, *e.g.*, postures could be defined as an attribute or a sub-event, the source of which could be replaced as soon as it keeps its label. Herein the posture status coming from MotionPod® and the Visual component are defined as attributes of the ontology model called “Person”, and the specific changes in these attribute values are modeled as primitive events.

Raw MotionPod® is post-processed using MotionPod® manufacturer software which provides the set of participant postures during the experiment (e.g., standing, sitting, lying down, changing from a posture to another). The vision component estimates standing and sitting postures based on a sitting threshold applied over the person 3D height. The sitting threshold is manually defined per video (and patient) based on the visualization of the participant average sitting height. A time synchronization step is performed based on sensors (video and accelerometer) readings timestamps.

The choice of which sensor posture estimation is used per activity model was based in the analysis of the system activity recognition performance in the defined contextual zones. For example, goal-oriented events that are performed farther from the camera viewpoint, where the Person remains immobile used accelerometer posture estimation, and Event models that requires quick response in posture estimation (Sitting to Standing transfer tests) were modeled in respect to the vision module posture estimation, as the accelerometer standard data acquisition rate is around 1 data acquisition/sec.

#### 4) *Events Models and IADLs*

Event models are described taking in account a person presence in a contextual zone, his/her proximity to a contextual objects, and his/her current body posture (*e.g.*, standing, bending). These constraints pose a limitation to some IADLs models as their target actions that cannot be directly detected (*e.g.*, writing a check to pay electricity bill, reading a magazine, organizing a set of cards) and they have the same contextual information.

In these cases, a few IADLs of the Clinical protocol are merged in a single activity. For example, answering and calling someone is a voice-based behavior in which the person interacts with an object (the phone), and, as there is none pressure sensor in the phone nor audio recognition in this approach, these two models are simplified to “using phone” activity. Table 4 lists the activity models of the clinical protocol (Scenario 02) and their simplified version.

TABLE 4. COMPARISON BETWEEN EVENT MODELS AND PROTOCOL IADLS

| EVENT MODEL | MODELED EVENT |
|---|---|
| Person using OfficeDesk | ▪ Make a shopping list of the ingredients for lunch<br>▪ Write a check to pay the electricity bill. |
| Person watching TV | ▪ Watch TV |
| Person preparing Coffee | ▪ Make Tea/Coffee. |
| Person using Pharmacy | ▪ Organize the prescribed drugs inside the drug box according weekly to the hour/day schedule |
| Person using Phone; | ▪ Call a taxi,<br>▪ Answer the Phone |
| Person watering Plant; | Watering the plant |
| Person using Reading Zone | Person reading the newspaper/magazine |

### D. *Evaluation*

Evaluation results are presented for event models based only on visual constraints and the Multi-sensor approach to evaluate the improvements in event detection performance. Vision-based models consist of the same activity models but only taking in account posture estimation obtained from visual information.

Activity recognition performance is assessed using the performance indices precision and sensitivity (Equations 1 and 2, respectively), where TP refers to True Positive events; FP to False Positive events, TN to True Negative events, and FN to False Negative events. The evaluation is performed for the 15 min. of each video sequence, and not in the form of previously extracted video chunks containing only IADL events.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

## III. RESULTS AND DISCUSSION

Table 5 shows the events detection performance rate for the vision system approach about 9 participants (15 min. each). GT states for ground-truth, and it specifies the total number of events inside the video sequences of all participants. The term “Person” was omitted from the event model name to improve the table readability. Table 6 shows the event detection rate for the multi-sensor approach.

TABLE 5. ACTIVITY DETECTION RATE OF THE VISION SYSTEM

| Activity | GT | Precision (%) | Sensitivity (%) |
|---|---|---|---|
| Standing and preparing Coffee | 13 | 37.5 | 37.50 |
| Sitting and using OfficeDesk | 8 | 5.80 | 100.00 |
| Standing and using OfficeDesk | 12 | 43.37 | 100.00 |
| Standing and using Pharmacy | 8 | 75.92 | 100.00 |
| Standing and using Phone | 19 | 83.33 | 25.92 |
| Standing and watering Plant | 10 | 100.00 | 100.00 |
| **Average performance** | 70 | 57.65 | 77.23 |

TABLE 6. ACTIVITY DETECTION RATE OF THE FUSION SYSTEM

| Activity | GT | Precision (%) | Sensitivit y (%) |
|---|---|---|---|
| Standing and preparing Coffee | 13 | 51.85 | 100 |
| Sitting and using OfficeDesk | 8 | 27.22 | 100 |
| Standing and using OfficeDesk | 12 | 43.37 | 100 |
| Standing and using Pharmacy | 8 | 75.92 | 100 |
| Standing and using Phone | 19 | 83.33 | 100 |
| Standing and watering Plant | 10 | 100.00 | 61.11 |
| **Average performance** | 70 | 63.61 | 93.51 |

The vision system has an average precision of 57.65% and a sensitivity of 77.23%. The multi-sensor system performance is higher at the two indices, 63.61% at precision and 93.51% at sensitivity, highlighting that the use of accelerometer data in posture estimation has increased the overall system performance by improving the precision of the first three events of Table 6.

Factors like occlusion, illumination changes, and color clothing similar to background have reduced person body detection at visual component, increasing the difficult at determining posture. Accelerometer data acquisition resolution is sometimes not sufficient to cope with person speed, e.g., in changing postures activities. But, the use of both sensors, each one applied in the situation (event model) where it provides better estimation, have improved the overall system performance without incorporating noisy in the event estimation framework.

Part of the FP events that caused lower values of precision could be explained by unexpected behaviors of the participants that are not predicted in the event models. For example, events based on the spatial zone "OfficeDesk" are affected by this zone proximity to the "Watching TV" zone. FP events of "Using pharmacy" events are explained by a few patients placing a chair into this zone to watch the TV (located in the opposite side of the room). These inconsistences will be future addressed by refining the event and scene models and by the adoption of new features (e.g., adding an orientation attribute to the "Person" model and place constraints avoiding the cited FP event).

## IV. Conclusion and Future Work

The preliminary results of the proposed system shows that the adoption of a multi-sensor approach (a video data complemented by a wearable accelerometer) has improved the event detection performance of the activity recognition system in comparison to the same system based only on video data.

Future work will focus at analyzing an unsupervised way of choosing among the posture attributes based on reliability measurement. A reliability measurement could support the automatic choice of one sensor over the others or the calculation of the likelihood of events based on multiple sensors reading. A broader validation is also planned to evaluate the reproducibility of results shown in a larger quantity of patients. The proposed multi-sensor approach also added new posture types (e.g., bending, lying down) that will be considered in the next evaluation. The long-term goal of the proposed approach is to support clinicians in the identification of emerging symptoms that could early diagnose Alzheimer patients at mild to moderate stages in a quantitative and objective way.

## V. References


[1] H. Maki, H. Ogawa, S. Matsuoka, Y. Yonezawa, and W. M. Caldwell, "A daily living activity remote monitoring system for solitary older people" Proc. Annual International Conference of the Engineering in Medicine and Biology Society (EMBC 11), Massachusetts 2011, pp. 5608–5611.

[2] Y. Yang and X. Zhao, "Development of a fall detection algorithm based on a tri-axial accelerometer," Proc. International Conference on Biomedical Engineering and Informatics (BMEI 11), vol. 3, 2011, pp. 1371–1374.

[3] H. Foroughi, B. S. Aski, and H. Pourreza, "Intelligent video surveillance for monitoring fall detection of older people in home environments," Proc. International Conference on Computer and Information Technology (ICCIT 2008), Issue Date: 24-27 Dec. 2008, pp. 219-224.

[4] A. H. Nasution, P. Zhang, and S. Emmanuel, "Video surveillance for older people monitoring and safety," Proc. IEEE Region 10 Conference (TENCON 2009), 2009, pp. 1–6.

[5] A. H. Nasution and S. Emmanuel, "Intelligent video surveillance for monitoring older people in home environments," Proc. IEEE 9th Workshop Multimedia Signal Processing, 2007, pp. 203–206.

[6] M.-L. Wang, C.-C. Huang, and H.-Y. Lin, An intelligent surveillance system based on an omnidirectional vision sensor," Proc. IEEE Conference on Cybernetics and Intelligent Systems, i. 7-9 June 2006, pp. 1–6.

[7] S. Fleck, and W. Strasser, "Smart camera based monitoring system and its application to assisted living," Proc. of the IEEE, vol. 96: 10, 2008, pp. 1698–1714.

[8] L. Monaci and R. G. Morris, "Neuropsychological screening performance and the association with activities of daily living and instrumental activities of daily living in dementia: baseline and 19- to 24-month follow-up", International Journal of Geriatric Psychiatry, vol. 27:2, 2012, pp. 197-204.

[9] O. Ojetola, E. I. Gaura, and J. Brusey, "Fall detection with wearable sensors--safe (Smart Fall Detection)", Proc. International Conference in Intelligent Environments (IE 2011), 25-28 July 2011, pp. 318–321.

[10] A. Fleury, M. Vacher, and N. Noury, "SVM-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results", IEEE Transactions in Information Technology in Biomedicine, March 2010, vol. 14 (2), pp. 274–283.

[11] A. Fleury, N. Noury, and M. Vacher, "Introducing knowledge in the process of supervised classification of activities of Daily Living in Health Smart Homes," Proc. e-Health Networking Applications and Services (Healthcom 2010), IEEE Press, 1-3 July 2010, pp. 322–329.

[12] S. Zhang, S. McClean, B. Scotney, X. Hong, C. Nugent, and M. Mulvenna, "Decision support for alzheimer's patients in smart homes," Proc. Computer-Based Medical Systems (CBMS '08), 17-19 June 2008, pp. 236–241.

[13] O. Brdiczka, M. Langet, J. Maisonnasse, and J. Crowley, "Detecting human behaviors models from multimodal observation in a smart home", IEEE Trans. on Automation Science and Engineering, vol. 6(4):1, 2009, pp. 588-597.

[14] I. Gómez-Conde, D.N. Olivieri, X.A. Vila, and L. Rodríguez-Liñares, "Smart telecare video monitoring for anomalous event detection," Proc. Iberian Conference on Information Systems and Technologies (CISTI 10), 2010, pp. 1–6.

[15] N. Zouba, F. Bremond, and M. Thonnat, "An activity monitoring system for real older people at home: validation study," Proc. IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 10), IEEE Press., 2010, pp. 278 – 285.

[16] S. Gillain, E. Warzee; F. Lekeu, V. Wojtasik, D. Maquet, J.-L. Croisier, E. Salmon, and J. Petermans, "The value of instrumental gait analysis in older people healthy, MCI or Alzheimer's disease subjects and a comparison with other clinical tests used in single and dual-task conditions," Annals of Physical and Rehabilitation Medicine, vol.52, 2009, pp. 453–474.

[17] J. Cancela, M. Pastorino, M. T. Arredondo, M. Pansera, L. Pastor-Sanz, F. Villagra, M.A. Pastor and A. P. Gonzalez, "Gait assessment in Parkinson's disease patients through a network of wearable accelerometers in unsupervised environments," Proc. Annual International Conference on of the Engineering in Medicine and Biology Society (EMBC 11), Massachusetts, 2011, pp. 2233-2236.

[18] H. H. Manap, N. M. Tahir, A. I. M. Yassin, and R. Abdullah, "Anomaly gait classification of parkinson disease based on ann," Annals of Physical and Rehabilitation Medicine, vol.52, 2009, pp. 579–587.

[19] R. Romdhane, E. Mulin, A. Derreumeaux, N. Zouba, J. Piano, L. Lee, I. Leroi, P. Mallea, R. David, M. Thonnat, F. Bremond, and Ph. Robert,

"Automatic video monitoring system for assessment of alzheimers disease symptoms," Journal of Nutrition, Health and Aging, September 2011.

[20] D. P. Chau, F. Bremond, and M. Thonnat, "A multi-feature tracking algorithm enabling adaptation to context variations," Proc. International Conference on Imaging for Crime Detection and Prevention, 2011.

[21] M. Nambu, K. Nakajima, M. Noshiro, and Toshiyo Tamura, "An algorithm for the automatic detection of health conditions," Proc. Engineering in Medicine and Biology Magazine, IEEE Press., July-Aug. 2005, v.24:4, pp. 38–42.

[22] B. Auvinet, P. Touzard, D. Chaleil, C. Touzard, A. Delafond, C. Foucher, and F. Multon, "Dual tasking and gait in people with mild cognitive impairment according to amnesic and non-amnesic subgroups, preliminary results," Annals of Physical and Rehabilitation Medicine, vol. 54, S1, pp. e87–e94, 2011.

[23] M.F. Folstein, L.M. Robins, and J.E. Helzer, "The mini-mental state examination," Arch Gen. Psychiatry, 1983, 40(7): p. 812.

[24] N. Zouba, F. Bremond, and M. Thonnat, "Multisensor fusion for monitoring elderly activities at home", Proc. IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 09), Genoa, Italy, September 2-4, 2009, pp.98-103.

[25] S. Zaidenberg, B. Boulay, F. Bremond, and M. Thonnat, "A generic framework for video understanding applied to group behavior recognition," (AVSS 2012), in press.

[26] A.T. Nghiem, F. Bremond and M. Thonnat, "Controlling background subtraction algorithms for robust object detection", Proc. 3rd International Conference on Imaging for Crime Detection and Prevention, ICDP 09, Kingston University, London, UK, December, 2009, pp. 1-6.

[27] H. Medjahed, *et al*, A pervasive multi-sensor data fusion for smart home healthcare monitoring. Proc. IEEE International Conference on Fuzzy Systems (FUZZ 11), 27-30 June 2011, pp. 1466-1473.

[28] Y. Cao, L. Tao, and G. Xu, "An event-driven context model in elderly health monitoring", Proc. Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing, 2009, pp. 120-124, doi:10.1109/UIC-ATC.2009.47.

# Hierarchical Human Activity Recognition Using GMM

Prateek Srivastava
Department of Electrical and Computer Engineering,
National University of Singapore (NUS), Singapore
prateek.k.srivastava@gmail.com

Wai-Choong Wong
Department of Electrical and Computer Engineering,
National University of Singapore (NUS), Singapore
elewwcl@nus.edu.sg

***Abstract*— In this paper, we propose a hierarchical human activity recognition system using Gaussian mixture models (GMMs) on continuous daily activities. The system recognizes the human activities by making use of tri-axial accelerometer and bi-axial gyroscope. We use different features such as mean, variance, root mean square, pitch, and roll for activity classification. Comparative performance assessments are carried out using the publicly available Wearable Action Recognition Dataset (WARD). The hierarchical recognition happens in two steps. First, the test data is classified into two broad clusters – static activity and dynamic activity. Second, the recognition is carried out within the identified class. For continuous activity recognition, our proposed system is able to achieve a recognition accuracy of 86.92% which is 2.63% above the baseline system. The new algorithm also provides more flexibility for better feature selection for different sets of activities.**

***Keywords-Human Activity recognition; wearable sensors; pattern recognition; Gaussian Mixture modeling (GMM).***

## I. INTRODUCTION

Due to the recent progress in ubiquitous and wearable computing, activity recognition has become a major contributor towards many monitoring and interaction applications. Human action/activity recognition is an emerging field of research. Physical activity can be defined as "any bodily movement produced by skeletal muscles that result in energy expenditure above the resting level" [1]. The objective of any human activity recognition system is to recognize any human activity using its observed sensor/visual data. Generally, human activities can be classified into two broad categories – static (which involves minimal movement of body parts such as standing or sitting) and dynamic (which involves some motion in the body parts such as boxing or walking).

A popular approach to activity recognition is based on the use of visual data [1, 2], which is high-dimensional and dense. The visual data can sometimes become intrusive and disruptive, and hence, it poses a challenge to personal privacy. Moreover, the vision-based activity recognition systems are very sensitive to ambient lighting conditions and occlusion. With the recent miniaturization of simple sensors such as accelerometers and gyroscopes, researchers have begun to adopt the use of these low-powered, unobtrusive sensors.

### *A. Motivation*

Due to the increase in the aging population, the world will soon have an increased number of aging baby boomers. As the existing and the future heath care sectors cannot effectively serve all the baby boomers, there is an increased demand for health monitoring and support of elderly-care units using assisted living systems. Consequently, the need for remote health care systems for patient monitoring is gradually growing; see Ibrahim et al. [3].

In this paper, an inertial sensor framework is used for human activity recognition because it provides an unobtrusive, low-powered and cost effective solution for many applications such as daily assisted living of elder care, virtual-real world interaction, sports training, and long term monitoring purpose. Longer term monitoring would reveal the subject's activity levels with respect to metabolic energy expenditure, associated with different activities such as walking, standing, etc. Medical professionals believe that one of the best ways to detect an emerging medical condition before it becomes critical is to look for changes in the activities of daily living (ADLs), instrumental ADLs (IADLs), and enhanced ADLs (EADLs); see Tapia et al. [4].

### *B. Objective*

In this work, our objective is to build a fast and accurate system for recognizing human activities. This system uses the data from the accelerometers and gyroscopes so as to recognize continuous activities.

### *C. Organization*

This is paper is organized as follows. Section II discusses some prior works in the field of human action/activity recognition. In Section III, we discuss the approach and the procedure of our algorithm. In Section IV, the hierarchical action recognition approach is introduced. In Section V, the results are presented and discussed. Finally, Section VI concludes the paper with some recommendations about future work.

## II. LITERATURE REVIEW

Recent developments in sensor technology have led to miniaturized inertial sensors – accelerometer and gyroscope.

These inertial sensors are widely used by the wearable activity recognition researchers. Bao et al [5] and Eric et al [6] have studied activity recognition using multiple sensors at different locations on the body. Thomas et al [7] have used the multimodal approach of activity recognition by combining motion sensors with ultrasonic sensors for continuous activity recognition. The application domains of activity recognition systems are diversity with examples such as Ernst et al [9] to recognize moves in martial arts, while David et al [10] built a mixed reality car parking game based on human computer interaction.

A number of activity recognition algorithms have been explored and these include:

1. Decision Trees: It finds a set of thresholds for a pattern-dependent sequence of features. Bao et al [5] used C4.5 decision tree classifier.
2. Nearest Neighbor (NN): It assigns patterns to the majority of class among k nearest neighbor using a performance optimized value of k. Ravi et al [8], Bao et al [5], and Maurer et al [11] used NN classifiers for activity recognition.
3. Naïve Bayes: It is a Bayes theorem based probabilistic classifier. Bao et al [5], and Maurer et al [11]
4. Support Vector Machine (SVM): Ravi et al [8] have used SVM classifiers.
5. Hidden Markov Model: one of the most popular statistical model for capturing temporal patterns in the data. It is extensively used in activity recognition. Yamato et al [1] used HMM to recognize different tennis strokes.
6. Gaussian Mixture model (GMMs): GMMs are parametric representations of any probability density function. Allen et al [12] used GMMs for transactional activities with an accuracy of 76.6 %. Ibrahim et al [3] recognize simple activities with small set of activities with 88.76 %. In this paper, Gaussian mixture models are also used for recognition.

Pattern recognition highly depends on the kinds of features used to model the patterns. Thus, features are very crucial for any recognition system. Some time domain features that have been commonly used include:

1. Mean: The mean value feature has been used by Ravi et al [8], and Bao et al [5].
2. Variance: The variance feature has been used by Ravi et al [8]
3. Root Mean Square (RMS): The RMS feature has been used by Maurer et al [11].

In the literature, researchers have tried different window lengths for activity recognition such as 6.7s in Bao et al [5], and 5.12s in Ravi et al [8]. In this paper, we use a window length of 1s is used to classify the activities. We have chosen a lower window length to facilitate real-time recognition.

## III. Approach and Procedure

### A. Dataset

To present a general comparison of results, a public dataset 'WARD: A Wearable Action Recognition Database' collected by Yang et al [13, 14] is used. In the WARD database, sensor data of different continuous activities have been collected. It contains data corresponding to 20 different subjects and 13 different activities. It also contains non-transient human actions. In order to sufficiently sample the continuous movement of a non-transient action, each subject performs one trial of an action for more than 10 seconds. The sensors are placed at different locations of the subject's body. Each sensor contains a 3-axis accelerometer and a 2-axis gyroscope. The locations of the sensors on the body are shown in Figure 1.

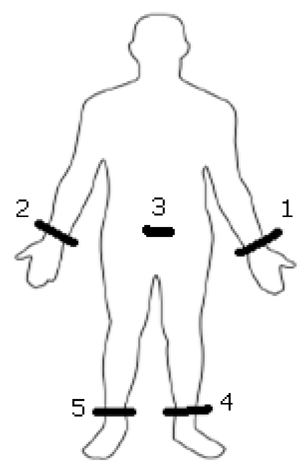



Figure 1. Sensor locations and orientation of a subject for WARD dataset, where the bold lines represent the sensor locations.

- Sensor 1: Outside center of the lower left forearm joint. The y-axis of the gyroscope points to the hand.
- Sensor 2: Outside center of the lower right forearm joint. The y-axis of the gyroscope points to the hand.
- Sensor 3: Front center of the waist. The x-axis of the gyroscope points down.
- Sensor 4: Outside center of the left ankle. The y-axis of the gyroscope points to the foot.
- Sensor 5: Outside center of the right ankle. The y-axis of the gyroscope points to the foot

This dataset contains the following 13 activities: 1. Rest at Standing (ReSt). 2. Rest at Sitting (ReSi). 3. Rest at Lying (ReLi). 4. Walk forward (WaFo). 5. Walk forward left-circle (WaLe). 6. Walk forward right-circle (WaRi). 7. Turn left (TuLe). 8. Turn right (TuRi). 9. Go upstairs (Up). 10. Go

 ISBN: 978-1-61208-235-6 

downstairs (Down). 11. Jog (Jog). 12. Jump (Jump) and 13. Push wheelchair (Push). For more details, please refer the WARD database manual [13].

### B. System Overview

Our human activity recognition system recognizes the activities using the WARD data. In this system, a statistical recognizer for different activities is built. The system is divided into two phases, namely the training phase and the testing phase. In the training phase, the raw data for all activities is first collected using the 3-axis accelerometer and 2-axis gyroscope. The sampled data from the accelerometer and gyroscope are combined. Then, suitable time domain features are extracted and are used to model the activities using Gaussian mixture models (GMMs). The models of all the different activities are stored at the end of the training phase. In the testing phase, activity data of the test subject is first collected and the features are extracted. Then, the maximum probability of match of the test sample against all stored sample patterns is calculated. That pattern, which has the highest likelihood of match against the test pattern, is recognized as the correct activity. The overall recognition system is depicted in Figure 2.

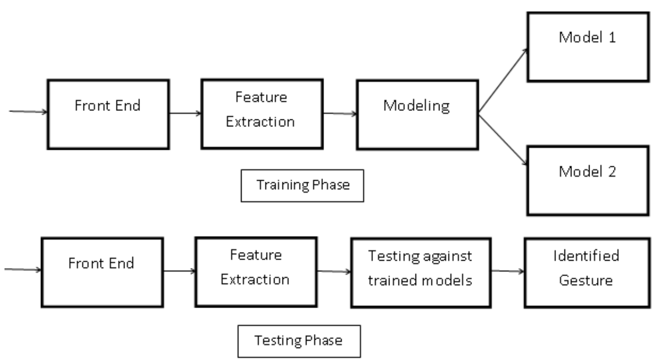



Figure 2. System overview of the activity recognition system.

### C. Features for activity recognition

We have used time domain features for performing activity recognition. This is because our initial set of experiments suggests that time domain features perform better as compared to frequency domain features. The time domain features that we have used for the classification are mean, variance and root mean square (RMS).

In order to represent the angular information, the captured accelerometer and gyroscope data are used to calculate the pitch and roll. Since we are using 3-axis accelerometer and 2-axis gyroscope, the pitch and roll are calculated using the following equations.

roll = arctan(-acc_y/ -acc_z);
pitch = arctan( acc_x, sqrt( acc_y * acc_y + acc_z * acc_z));

where, acc_x, acc_y and acc_z are the x-axis, y-axis and z-axis accelerometer values respectively.

### D. Activity Modeling using the GMM Algorithm

GMMs [15-16] are parametric representations of a probability density function. When trained to represent the distribution of a feature vector, GMMs can be used as classifiers. GMMs have proved to be a powerful tool for distinguishing time series data with different general properties. The use of GMMs for modeling activity is motivated by the interpretation that the (1) uni-variate Gaussian densities have a simple and concise representation, depending uniquely on two parameters, mean and variance, (2) they are capable of modeling arbitrary densities, (3) the Gaussian mixture distribution is universally studied and its behaviors are widely known, (4) a linear combination of Gaussian basis functions is capable of modeling a large class of sample distributions. In principle, the GMM can approximate any probability density function to an arbitrary accuracy.

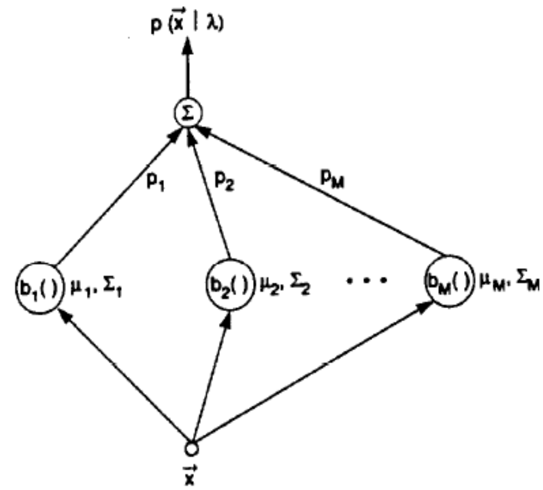



Figure 3. Depiction of an *M* component Gaussian mixture density [15].

A GMM is a weighted sum of M component densities as shown in Figure 3, given by the following equation:-

$$P\ (x_t\,|\lambda_s) = \sum_{i=1}^{M} p_i\, b_i\ (x(t))$$

Here, $x_t$ is a sequence of feature vectors from the activity data, $x(t)$ is a feature vector having $D$-dimensions. $b_i(s)$ is the Gaussian probability distribution function (PDF) associated with the $i^{th}$ mixture component and is given by:

$$b_i(x_t) = \frac{1}{2\pi^{D/2}\,|\Sigma_s^i|^{1/2}}\, e^{\left(-\frac{1}{2}\right)(x-\mu i)^{\wedge}T\, {\Sigma_i^s}^{-1}(x-\mu i)}$$

Here, $\mu_i$ is the mean vector and $\sum_i^s$ is the covariance matrix of the $i^{th}$ mixture component.

The mixture weights are such that :

$$\sum_{i=1}^{M} p_i = 1$$

Each trained activity is thus, represented by a Gaussian mixture model, collectively represented by

$$\lambda_s = \{\mu_i , \textstyle\sum_i^s , p_i \}$$

where $i = 1,2 ,\ldots M$, and $\mu_i$ , $\sum_i^s$ , $p_i$ represent the mean, covariance and weights of the $i^{th}$ mixture respectively.

## III. HIERARCHICAL RECOGNITION OF HUMAN ACTIVITY

A hierarchical recognition approach is proposed for human activity recognition on continuous activities. The dataset has data corresponding to two different types of activities – static and dynamic. In static activities such as sitting, standing, and lying, the motion sensor values are expected to have minimal variations over time. Apart from sitting, standing and lying, all other activities are categorized as dynamic since some motion is involved.

In our initial set of experiments, we observed that a number of static activities were mis-classified as dynamic activities and vice versa. In order to minimize such confusion, we propose to adopt a simple but powerful hierarchical approach to classify the full activity recognition on the continuous dataset. The first stage of this hierarchical approach tries to classify the test data into either static or dynamic activity categories. The second stage then tries to identify the correct activity from the chosen activity category. Together with improving the overall recognition accuracy, this approach also speeds up the process of computation because after the test activity is classified as static or dynamic in the first step, it will then, only be compared against its activity class for final recognition. The recognition flow of this algorithm is as shown in Figure 4.

This algorithm provides two major advantages over the baseline one-step classification system. First, since the two activity clusters: static and dynamic differ a lot in their corresponding statistical properties, the accuracy of classifying into static or dynamic classes is high. Since the test data is only compared with activities of its identified class in the second stage, the possibility of inter-class misclassification is eliminated. Second, this hierarchical system provides greater flexibility to extract different features for different clusters so as to improve the overall system performance.

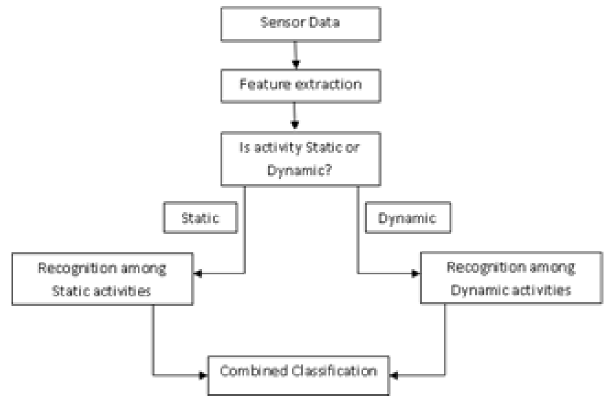



Figure 4. Flow chart of Hierarchical human activity recognition.

## IV. RESULTS AND DISCUSSION

### *A. Continuous Activity Recognition (CAR) on WARD dataset*

Different sets of experiments are performed to fine-tune the accuracy and speed of the recognition system. To achieve low latency, we use a window length of 1s in our recognition algorithm. The feature set includes pitch, roll, mean and variance from the accelerometer and gyroscope sensor data. To get the best number of mixture models, different numbers of mixture models are used to test the overall performance. The system with 32 mixture models gave the best overall performance, and thus we choose to model each activity using 32-GMMs. The results of the one-tier recognition system show an overall accuracy of 84.69 %. The confusion matrix of this system is shown in Table II.

From the confusion matrix of the one tier system it can be observed that a large number of static activities is misclassified as dynamic activities. For example, the misclassification accuracy of Rest at Sitting (Static activity) as Push Wheelchair (Dynamic Activity) is 12.2%.

A hierarchical algorithm using a two-step recognition process is proposed to classify the human continuous activities. The hierarchical recognition is proposed to minimize such inter-class misclassification as discussed above. To compare the results of the hierarchical algorithm over the one tier algorithm, we have used the same set of features {pitch & roll along mean and variance from accelerometer and gyroscope sensor data} and modeled them using 32-GMMs. The results of the cluster classification in the first step are shown in Table I. The overall accuracy of cluster classification is 96.58% where it is weighted against the number of test files for static and dynamic activities.

 ISBN: 978-1-61208-235-6

TABLE I. CLUSTER ACCURACY OF HIERARCHICAL RECOGNITION SYSTEM.

| | Static | Dynamic |
|---|---|---|
| Static | 95.5 | 4.5 |
| Dynamic | 3.0 | 97.0 |

In Table II, the accuracy of the one tier system with the same feature set and the same number of Gaussian mixtures as is used in the hierarchical system is 84.69 %. The accuracies of the hierarchical activity recognition system are shown in Table III. Here, we can observe that the overall accuracy of the hierarchical activity recognition system is 86.92%. This shows that the hierarchical algorithm improves the system performance by 2.63 % over the one tier system. It can also be observed from Table 3 that the misclassification of static activities as dynamic activities and vice versa is reduced. For example, the misclassification accuracy of Rest at Sitting (Static activity) as Push Wheelchair (Dynamic Activity) is 5.21% as compared to 12.2% in one tied system.

## V. CONCLUSION AND RECOMMENDATIONS

In this paper, we have proposed a hierarchical human activity recognition system using Gaussian mixture models (GMMs). The results of the system are competitive as compared to prior activity recognition systems. The performance of the proposed system is tested using the publicly available WARD dataset to provide a better comparability. An overall system accuracy of 86.92 % is achieved using this hierarchical approach, which is an improvement of 2.63% over the baseline system. The proposed hierarchical algorithm also provides the flexibility to use different feature sets for the identification of different classes of activities. Since the static activities and dynamic activities differ a lot in their statistical properties, the best performing feature sets for these clusters can be used to obtain the best performance in each individual cluster.

Another important contribution of this paper is that the recognition is performed using less test data. In continuous activities, the activities are recognized using 1 sec of test data. Also, to capture the angular information, the pitch and roll using the accelerometer and gyroscope are used as the feature set. To further reduce the time of recognition, the concept of fast elimination of such patterns that certainly do not match a given behavioral pattern can be adopted. This concept has been proposed by Bajan [17].

In this paper, the sensors are placed at five different locations so as to capture full body motion statistics. The sensor placement is maintained as per the WARD Dataset so as to maintain the comparability of results. The orientation of sensors may not be optimum as this orientation of the sensors does not exploit the symmetry of body. In future, there is need to find the optimal number, orientation and placement of sensors required to perform human activity recognition.

TABLE II. CONFUSION MATRIX OF ONE TIER ACTIVITY RECOGNITION SYSTEM (IN PERCENTAGE).

| | **ReSt** | **ReSi** | **ReLi** | **WaFo** | **WaLe** | **WaRi** | **TuLe** | **TuRi** | **Up** | **Down** | **Jog** | **Jump** | **Push** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **ReSt** | 68.2 | 11.2 | 0 | 0 | 0.5 | 1.8 | 0 | 0 | 5.8 | 6.3 | 0 | 0 | 6.3 |
| **ReSi** | 0.5 | 74.7 | 0 | 0 | 0 | 0 | 0 | 0 | 4.5 | 4.1 | 0 | 4.1 | 12.2 |
| **ReLi** | 0 | 4.6 | 75.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 19.8 | 0 |
| **WaFo** | 0 | 0 | 0 | 84.8 | 3.2 | 3.8 | 0 | 0 | 1.3 | 1.9 | 0.6 | 4.4 | 0 |
| **WaLe** | 0 | 0 | 0 | 1.9 | 93.7 | 0 | 3.4 | 0 | 0 | 0 | 0 | 0.5 | 0.5 |
| **WaRi** | 0 | 0 | 0 | 2.5 | 0 | 86.4 | 3 | 2 | 1 | 2 | 2.5 | 0.5 | 0 |
| **TuLe** | 0 | 0 | 0 | 0 | 11.1 | 0 | 88.9 | 0 | 0 | 0 | 0 | 0 | 0 |
| **TuRi** | 0 | 0 | 0 | 1 | 0 | 5.6 | 0 | 88.3 | 0 | 3.1 | 1.5 | 0.5 | 0 |
| **Up** | 0 | 0 | 0 | 1.6 | 1.1 | 1.1 | 0 | 0 | 65.2 | 19.3 | 9.6 | 0 | 2.1 |
| **Down** | 0 | 0 | 0 | 0 | 1.2 | 0 | 0 | 0 | 1.8 | 84.4 | 10.8 | 1.2 | 0.6 |
| **Jog** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2.9 | 0 | 97.1 | 0 | 0 |
| **Jump** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6 | 0.6 | 98.8 | 0 |
| **Push** | 0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6 | 0.6 | 0 | 3.5 | 94.8 |

TABLE III. CONFUSION MATRIX OF OVERALL SYSTEM ACCURACY OF HIERARCHICAL RECOGNITION SYSTEM (IN PERCENTAGE)

| | ReSt | ReSi | ReLi | WaFo | WaLe | WaRi | TuLe | TuRi | Up | Down | Jog | Jump | Push |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **ReSt** | 70.4 | 11.2 | 0 | 0 | 0 | 3.6 | 4 | 0.9 | 0 | 6.3 | 0 | 0 | 3.6 |
| **ReSi** | 6.6 | 83.9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4.3 | 0 | 0 | 5.2 |
| **ReLi** | 0 | 5.7 | 94.3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **WaFo** | 0 | 0 | 0 | 84.2 | 3.2 | 3.8 | 0 | 0 | 1.3 | 1.9 | 0.6 | 4.4 | 0 |
| **WaLe** | 0 | 0 | 0 | 1.9 | 93.7 | 0 | 3.4 | 0 | 0 | 0 | 0 | 0.5 | 0.5 |
| **WaRi** | 0 | 0 | 0 | 2.5 | 0 | 87.3 | 3.1 | 1.5 | 1 | 2 | 2 | 0.5 | 0 |
| **TuLe** | 0 | 0 | 0 | 0 | 11.1 | 0 | 88.9 | 0 | 0 | 0 | 0 | 0 | 0 |
| **TuRi** | 0 | 0 | 0 | 1 | 0 | 5.6 | 0 | 88.8 | 0 | 3.1 | 1 | 0.5 | 0 |
| **Up** | 0 | 0 | 0 | 1.6 | 1.1 | 0.5 | 0 | 0 | 66.3 | 18.5 | 9.8 | 0 | 2.2 |
| **Down** | 0 | 0 | 0 | 0 | 1.2 | 0 | 0 | 0 | 1.8 | 84.4 | 10.8 | 1.2 | 0.6 |
| **Jog** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2.9 | 0 | 97.1 | 0 | 0 |
| **Jump** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6 | 0.6 | 98.8 | 0 |
| **Push** | 1.2 | 4.1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2.9 | 91.8 |

## REFERENCES


[1] J. Yamato, J. Ohya, and K. Ishii, “Recognizing human action in time-sequential images using hidden markov model”, International IEEE Conference on Computer Vision and Pattern Recognition (CVPR ‘92), 1992, pp. 379–385.

[2] T. Starner, “Visual recognition of american sign language using hidden markov models”, Master’s thesis, Massachusetts Institute of Technology, Cambridge, MA, USA (Feb 1995).

[3] R. Ibrahim, E. Ambikairajah, B. Celler, and N.H. Lovell, “Time-Frequency based features for classification of walking patterns”,15th Int. Conf. on Digital Signal Processing (DSP ’07), 1-4 Jul 2007, pp. 187–190.

[4] E. M. Tapia, S. S. Intille, and K. Larson, “Activity Recognition in the Home Using Simple and Ubiquitous Sensors”, Pervasive Computing, LNCS, Springer, 2004, Vol. 3001/2004, pp. 158-175.

[5] L. Bao and S. Intille, “Activity recognition from user-annotated acceleration data”, Pervasive Computing, LNCS, Springer, 2004, Vol. 3001/2004, pp. 1-17.

[6] E. Guenterberg, H. Ghasemzadeh, V. Loseu, and R. Jafari, “Distributed continuous action recognition using a hidden markov model in body sensor networks”, Distributed Computing in Sensor Systems, LNCS, Springer, 2009, Vol. 5516/2009, pp. 145-158.

[7] T. Stiefmeier, G. Ogris, H. Junker, P. Lukowicz, and G. Tröster, “Combining motion sensors and ultrasonic hands tracking for continuous activity recognition in a maintenance scenario”, 10th Int. Symp. on Wearable Computing (ISWC ’06), 11-14 Oct 2006, Zurich, Switzerland, pp. 97-104.

[8] N. Ravi, N. Dandekar, P. Mysore, and M. Littman, “Activity recognition from accelerometer data”, IAAI'05 Proceedings of the 17th conference on Innovative applications of artificial intelligence - Volume 3 pp 1541-1546.

[9] E. A. Heinz, K. S. Kunze, M. Gruber, D. Bannach, and P. Lukowicz, “Using wearable sensors for real-time recognition tasks in games of martial arts – An initial experiment”, 2nd IEEE Symp. on Computational Intelligence and Games (CIG ‘06), 22-24 May 2006, Reno, NV, USA, pp. 98–102.

[10] D. Bannach, O. Amft, K.S. Kunze, E.A. Heinz, G. Tröster, and P. Lukowicz, “Waving real hand gestures recorded by wearable motion sensors to a virtual car and driver in a mixed-reality parking game”, IEEE Symp. on Computational Intelligence and Games (CIG ‘07), 1-5 Apr 2007, Colchester, U.K., pp. 32–39.

[11] U. Maurer, A. Smailagic, D. Siewiorek, and M. Deisher, “Activity recognition and monitoring using multiple sensors on different body positions”, BSN '06 Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks pp 113 - 116.

[12] F. Allen, E. Ambikairajah, N. Lovell, and B. Celler, “An adapted gaussian mixture model approach to Accelerometry-Based movement classification using Time-Domain features”, 28th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBS ’06), 30 Aug – 3 Sep 2006, New York, NY, USA, pp. 3600–3603.

[13] A. Yang, P. Kuryloski, and R. Bajcsy, “WARD: A Wearable Action Recognition Database”, 27th Annual CHI Conference, 4-9 Apr 2009, Boston, MA, USA.

[14] A. Yang, R. Jafari, S. Sastry, and R. Bajcsy, “Distributed recognition of human actions using wearble motion sensor networks”, Jour. of Amibient Intelligence and Sensor Environments, Vol. 1, no. 2, 2009, pp. 103-115.

[15] D. A. Reynolds, “A Gaussian mixture modeling approach to text independent speaker identification”, Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA, USA, September 1992.

[16] S. Roberts, D. Husmeier, I. Rezek, and W. Penny, “Bayesian approaches to gaussian mixture modeling”, IEEE Trans. On Pattern Anal. Machine Intell., vol. 20, no. 11, Nov. 1998, pp. 1133–1142.

[17] J. Bazan, “Hierarchical classifiers for complex spatio-temporal concepts”, Trans. On Rough Sets IX, LNCS, Springer, 2008, Vol. 5390/2008, pp. 474-750.

# Autonomous Wheelchair: Concept and Exploration

Rafael Berkvens, Wouter Rymenants, Maarten Weyn, Simon Sleutel, and Willy Loockx
*Dept. of Applied Engineering: Electronics-ICT*
*Artesis University College Antwerp*
*Antwerpen, Belgium*
{*rafael.berkvens, wouter.rymenants*}*@student.artesis.be*
{*maarten.weyn, simon.sleutel, willy.loockx*}*@artesis.be*

***Abstract*—This paper is part of a larger project with as goal the construction of an autonomous wheelchair. We implemented some existing algorithms and made the first hardware setup for the project. We found that a lot of optimization is still necessary, but could indicate in what direction this research should first be aimed. For our localization approach, using FastSLAM, landmark data association is key. We use Vector Field Histograms as obstacle avoidance planning algorithm, which still has problems with local extrema. With these problems will be dealt in future research.**

***Keywords-Autonomous wheelchair; sensor fusion; FastSLAM; curvature scale space; obstacle avoidance.***

## I. Introduction

Autonomous vehicles have had major breakthroughs in recent years. This is shown in, among others, the DARPA Grand Challenge [1], [2], the autonomous vehicle from Stanford University [3], and Google's self-driving car [4]. There are a lot of problems involved in building a vehicle that can fully autonomously drive through an urban environment. Most crucial in the process are the planning, the control, and the perception of the vehicle [1]–[4].

This paper is part of a larger project with as goal the construction of an autonomous wheelchair. The project starts with a simple electrical wheelchair and attempts to equip this with various sensors and actuators. These sensors and actuators are then intelligently managed by a computer program, such as happens in regular autonomous vehicles.

Other research on autonomous wheelchairs, such as the ones described by Yanco [5], has been focused on navigating in small environments, assisted driving, and human-AI interface. In contrast, this project is focused on opportunistic outdoor navigation and obstacle avoidance. Concretely this means the wheelchair will use detailed information about the environment if it is available, but can also work without this information. This should not influence anything but navigation performance.

The goal of this research is to make a first implementation of the wheelchair, both in hardware and software. This is done by assembling the hardware and getting it operational as well as implementing the intelligence framework. The intelligence framework is a collection of existing algorithms that are commonly used in autonomous vehicles to perform the steering, localization, and planning.

In Section II, we describe how we implemented these existing algorithms. First is the obstacle avoidance in Section II-A, using the Vector Field Histogram (VFH) [6]. Second is the localization in Section II-B, using landmarks extracted from a laser range measurement in the FastSLAM [7] algorithm. The last algorithm, in Section II-C, is an implementation of A* search [8] to perform planning. Section III gives the performance of the obstacle avoidance and localization algorithms. These tests do not represent the overall performance of the algorithms. The main goal of the paper was collecting the algorithms and find out where optimization is needed. In Section IV, we conclude by describing how the tested algorithms must be extended or altered to optimize their performance in our application.

## II. Methods

We developed a program that cycles through three modules: localization, steering, and navigation. The navigation module performs a routing algorithm and decides which path to take on a global level. The steering module attempts to follow parts of the route at a time, dictated by the navigation module on a local level. In addition, the wheelchair will also avoid obstacles detected by the laser. The program accomplishes moving and avoiding, by sending commands to an IC that translates the high level commands into amounts of power to each motor. After the wheelchair performs the commands by the mobilization module, the localization module then calculates the location of the wheelchair. The hardware setup is shown in Figure 1.

### *A. Obstacle Avoidance*

Since wheelchairs are often used in an urban environment, avoiding obstacles is an important issue. The problem, however, is complex as obstacles have to be detected and the distance to the object needs to be calculated. Moreover the program has to identify whether the wheelchair is on a collision course. Moving objects further complicate obstacle avoidance.

How complex the obstacle avoidance algorithm needs to be depends mostly on the application. For a primitive

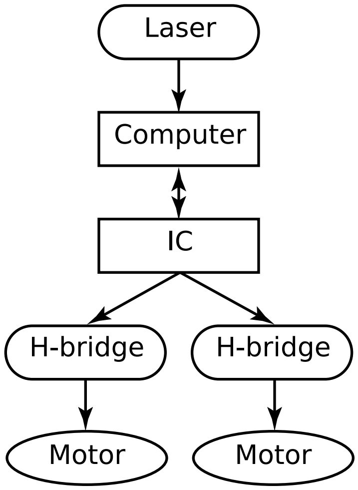



Figure 1. Wheelchair hardware setup

wheelchair, it could be sufficient to come to a full stop when an obstacle is detected. More sophisticated algorithms can decide on the position and size of an object, and plan a route around it (e.g., edge detection methods [9]). The set-back is that long computation times are necessary. In addition implementation is often complex. For some algorithms it is even necessary to stop the wheelchair to calculate the alternate route, before the wheelchair can resume its activities.

The solution utilized in this paper is the Vector Field Histogram (VFH) [6], which is an improvement on the Virtual Force Field (VFF), of the same author [10]. This method utilizes the concept of imaginary forces that influence the entity, in our case a wheelchair. This idea was first introduced by Khatib [11], and it lends itself very well to an effective way of avoiding obstacles. Any obstacle that is detected, will emit a force towards the wheelchair that repels it, the closer the object, and the higher the certainty of their actually being an object there, the stronger the force. On the other hand there is a goal, which attracts the wheelchair towards it's position. The combination of all these forces will decide the bearing of the wheelchair.

*1) Vector Field Histogram:* VFH implements this in three steps: Creating a belief state in the form of a 2D cartesian Histogram, a data reduction by transforming the 2D histogram to a 1D polar histogram, and finally using this data to create a new bearing for the wheelchair.

In the 2D Cartesian Histogram representation, each object in the world is mapped in a 2D grid. Each cell represents a part of the space, and holds the probability of there being an obstacle in that position. Whenever a new sample of sensor data suggests an obstacle in a certain cell, the probability of there being an obstacle increases. The wheelchair and the destination also hold a position in this grid, the 2D Cartesian histogram is an entire overview of the area of operation. This grid is also implemented in the VFF algorithm. A problem with using just the information on this grid to map to a wheelchair's bearing lies in the discrete nature of the grid. When the wheelchair moves to another cell on the grid, the force vectors will change direction and magnitude very sudden, causing fluctuations in the steering. Smoothing is wanted, but also slows down the algorithms reaction to sudden changes to the world. The solution proposed in the VFH algorithm is doing a second data reduction to a 1D Polar histogram, and smoothing this histogram instead. This will lower the impact of smoothing on the wheelchair's reaction to sudden changes, but will solve the fluctuations in the steering.

The polar histogram contains information on the the obstacle density in a certain direction (sector). Sectors with a low object density are called valleys, while sectors with a high object density are called peaks. This density is smoothed using a moving average filter, because of the data reduction this smoothing will have a lot less impact than doing it on the 2D Cartesian Histogram. Valleys are directions the wheelchair should consider taking, where as peaks are directions that are too cluttered with objects. Consequently it can be said that peaks repel the wheelchair, and valleys attract it. Which direction will be chosen further depends on the proximity of valleys to the sector of the destination. This is illustrated in Figure 2. The dashed fields in the circle are peaks, and the wheelchair should avoid moving towards them. The empty fields are valleys, which are available routes that should be considered. The dot in the upper right corner is the destination. Which valley is the most optimal is finally decided by their proximity to the object. The bearing that differs the least from the direct path will be best suited. However, to avoid paths that are too narrow for the wheelchair, the path that is finally chosen will not be the optimal path, but the optimal path + a constant. On the figure, the lines towards the wall in front of the destination would be the optimal path, where the arrows would be the actual chosen paths.

### *B. Localization*

To calculate the wheelchair's new position after a movement, a laser rangefinder is used. In an environment of which no prior knowledge is available to the wheelchair, this problem is similar to the Simultaneous Localization and Mapping (SLAM) [12] problem.

A map of the environment is assumed to be available in our setup. Using this assumption it is possible to perform SLAM on abstract, natural landmarks that are easily matched in subsequent observations of the environment. It is harder to identify these abstract landmarks as parts of the environment such as wall corners. This is not of concern here, since no map of the environment must be constructed.

The localization thus happens by finding landmarks in the environment. These landmarks are stored as their position

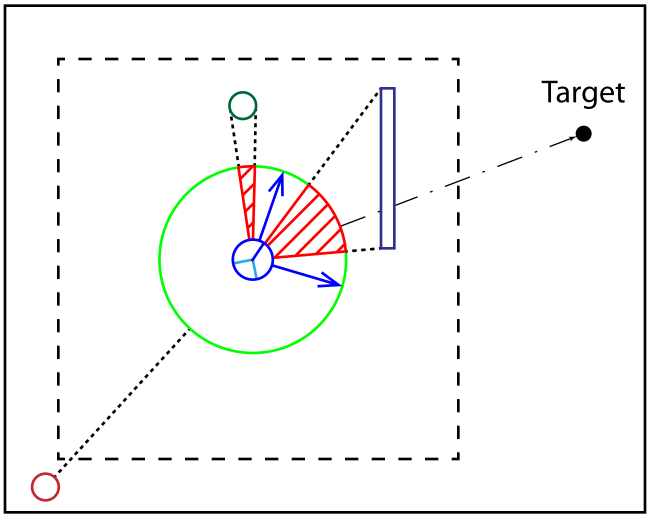



Figure 2. Vector Field Histogram imaginary forces and decisions.

in the environment, calculated from their position relative to the wheelchair. The movement of the wheelchair is then applied reversely to the landmarks, simulating wheelchair movement. When at the new position landmarks are detected, they are matched with the simulated location. These steps are visible in Figure 3.

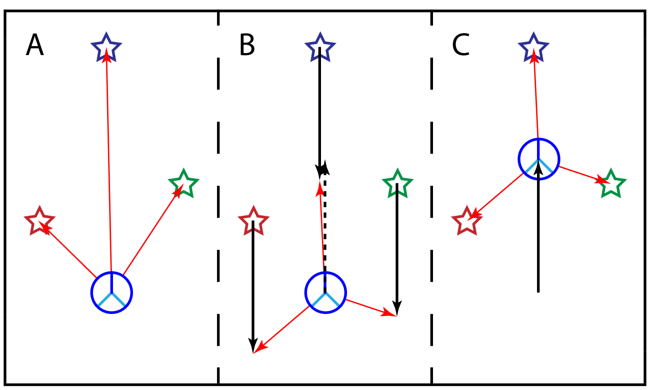



Figure 3. Localization using landmarks. A: Find landmarks, calculate their position in the environment. B: Simulate movement on landmarks, calculate position to wheelchair. C: Check found landmarks at new position with simulated locations of landmarks.

Our wheelchair searches for landmarks in the environment using a SICK LMS100 [13] laser rangefinder. Figure 4 shows a sample measurement of this device. It has a 18 meter range and operating angle of 270 degrees, scanning at every half degree [14].

To find distinctive landmarks in the range scan an algorithm based on the curvature scale space is used [15], [16]. The range scan is first parameterized to a curve based on the path length parameter. This curve is then convoluted with a Gaussian of varying scales. The convolution with the wider Gaussian is used to identify good landmarks, the convolution with the smaller Gaussian is used to localize a landmark. The complete algorithm is described below.

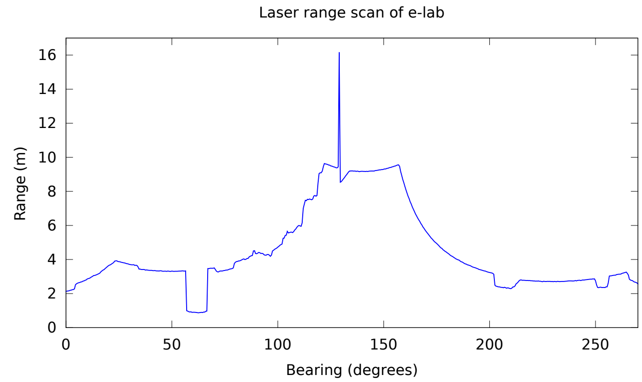



Figure 4. Measurement of the LMS100 laser rangefinder. The horizontal axis depicts the bearing of the measurement, the vertical axis depicts the range at that angle.

*1) Landmark Extraction:* The curvature scale space representation of the laser range scan is obtained by Equation (1). For a complete description of how the curvature $\kappa(s,\sigma)$ is derived, see [15].

$$\kappa(s,\sigma) = \dot{X}(s,\sigma)\ddot{Y}(s,\sigma) - \ddot{X}(s,\sigma)\dot{Y}(s,\sigma), \qquad (1)$$

where:

$$\dot{X}(s,\sigma) = x(s) \otimes \dot{g}(s,\sigma) \qquad (2)$$

$$\ddot{Y}(s,\sigma) = y(s) \otimes \ddot{g}(s,\sigma) \qquad (3)$$

$$\ddot{X}(s,\sigma) = x(s) \otimes \ddot{g}(s,\sigma) \qquad (4)$$

$$\dot{Y}(s,\sigma) = y(s) \otimes \dot{g}(s,\sigma) \qquad (5)$$

$$g(s,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-s^2/2\sigma^2} \qquad (6)$$

$$\dot{g}(s,\sigma) \equiv \frac{\partial g(s,\sigma)}{\partial s} = \frac{-1}{\sigma^3\sqrt{2\pi}} s e^{-s^2/2\sigma^2} \qquad (7)$$

$$\ddot{g}(s,\sigma) \equiv \frac{\partial^2 g(s,\sigma)}{\partial s^2} = \frac{1}{\sigma^3\sqrt{2\pi}} \left( \left[\frac{s}{\sigma}\right]^2 - 1 \right) e^{-s^2/2\sigma^2} \qquad (8)$$

The curve $x(s)$, used in (2) and (4), is a parameterization of the bearing of the range scan, see the dashed curve in Figure 5. The curve $y(s)$, used in (3) and (5), is a parameterization of the range of the range scan, see the solid curve in Figure 5.

The parameterization is done by creating the path length parameter $s$. The parameter depends on the distance between two subsequent data points, normalized from 0 to 1. Two new curves are constructed depending on this parameter as such: $\{(x_i, s_i), (y_i, s_i); i = 0, 1, \ldots, m\}$ with $s_0 = 0, s_m = 1$.

These curves are convoluted with derivatives of a Guassian $g(s,\sigma)$, see Equation (6). The first derivative of this

 ISBN: 978-1-61208-235-6 

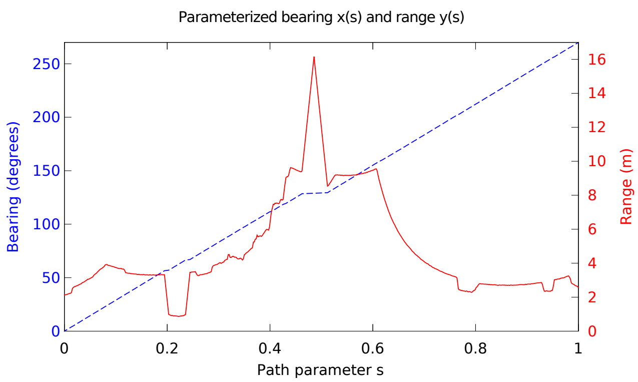



Figure 5. Parameterized curves of bearing and range. The left vertical axis depicts the bearing of the measurement, the right vertical axis depicts the range of the measurement. The horizontal axis depicts the path length parameter $s$.

Gaussian, Equation (7), is needed in (2) and (5). The first fifty points of this Gaussian are given as the dashed curve in Figure 6. The second derivative of this Gaussian, Equation (8), is needed in (4) and (3). The first fifty points of this Gaussian are given as the solid curve in Figure 6. Other points of these curves are all negligible close to zero. Since the Gaussians depend on the path length parameter $s$, which ranges from 0 to 1, they only exist over the positive side of the horizontal axis.

These Gaussians are scaled by varying $\sigma$. The scales range from $\sigma = 5 \times 10^{-4}$ to $\sigma = 1.05 \times 10^{-2}$ with a step width depending on the desired number of scales. Since four convolutions are needed to calculate the curvature of one scale, the number of scales is, however heuristically chosen in our implementation, a non-trivial number to choose.

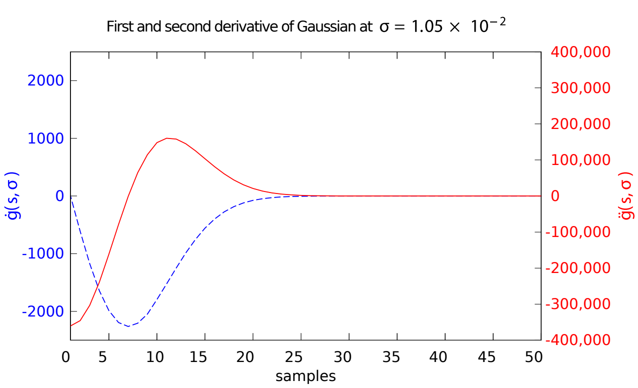



Figure 6. Gaussian derivatives with which the parameterized curves are convoluted. The left vertical axis depicts the first derivative of the Gaussian, the right vertical axis depicts the second derivative of the Gaussian.

Local extrema at larger scales are used to select landmarks that are expected to be easily identified again in subsequent scans. Local extrema at smaller scales are used to localize these landmarks [17]. Thus, these local extrema at both ends must be matched to each other. This matching is visible in Figure 7 as the lines. These lines range from the bottom, were the wider scales are and landmarks are selected, to the top where they are localized.

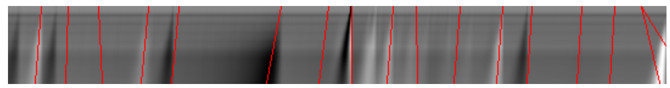


Figure 7. Curvature scale space of laser range scan with 64 scales. At the top of the figure are the curvatures with smaller scale, at the bottom are the curvatures with wider scale.

After selection and localization, the landmarks are used in a SLAM algorithm.

*2) FastSLAM:* FastSLAM is one modern algorithm to perform SLAM using landmarks [7], [18]–[20]. A landmark is in fact a feature of the environment saved as the range and bearing to that feature. By predicting the location of the landmarks after movement and matching this to the actual measurement, it is possible to select a most probable explanation.

The FastSLAM posterior is defined as follows [7]:

$$p(s^t, \theta | z^t, u^t, n^t) = p(s^t | z^t, u^t, n^t) \prod_{n=1}^{N} p(\theta_n | s^t, z^t, u^t, n^t) \tag{9}$$

where:

| | |
|---|---|
| $s^t$ | complete path of the wheelchair |
| $\theta$ | set of all $n$ landmark positions |
| $z^t$ | set of all observations |
| $u^t$ | set of all controls |
| $n^t$ | set of all data associations |
| $N$ | total number of landmarks |

The wheelchair path posterior, the first term in (9), is computed using a particle filter in order to pursue multiple possible paths and select the most probable. For each particle a set of $N$ Extended Kalman Filters (EKFs) is maintained to calculate the landmark position posteriors, the second term in (9).

The control in our system is expressed as a traveled distance and turn, thus the motion model incorporates distances instead of speeds:

$$s'_{t,x} \sim s_{t-1,x} + \cos(s_{t-1,\theta}) N(u_{t,d}, \alpha_1 u_{t,d}) \tag{10}$$
$$s'_{t,y} \sim s_{t-1,y} + \sin(s_{t-1,\theta}) N(u_{t,d}, \alpha_1 u_{t,d}) \tag{11}$$
$$s'_{t,\theta} \sim s_{t-1,\theta} + N(u_{t,\theta}, \alpha_2 u_{t,\theta}) \tag{12}$$

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-235-6 

where:

| | |
|---|---|
| $s_t'$ | predicted current state of the wheelchair |
| $s_{t-1}$ | previous state of the wheelchair |
| $u_{t,d}$ | distance in current control |
| $u_{t,\theta}$ | turn in current control |

Equations (10) and (11) predict a new location of the wheelchair after moving a certain distance $u_{t,d}$. Equation (12) predicts the new heading of the wheelchair after turning a certain angle $u_{t,\theta}$. In this model, it is assumed that the wheelchair first drives the distance and then turns in place. The variance of the error is captured in the $\alpha$ parameters.

Landmarks are then incorporated using EKFs. The measurement function $g(s_t, \theta_{n_t})$ is defined as:

$$g(s_t, \theta_{n_t}) = \begin{bmatrix} r(s_t, \theta_{n_t}) \\ \phi(s_t, \theta_{n_t}) \end{bmatrix} = \begin{bmatrix} \sqrt{(\theta_{n_t,x} - s_{t,x})^2 + (\theta_{n_t,y} - s_{t,y})^2} \\ \tan^{-1}1\left(\frac{\theta_{n_t,y} - s_{t,y}}{\theta_{n_t,x} - s_{t,x}}\right) - s_{t,\theta} \end{bmatrix} \quad (13)$$

The range $r$ and bearing $\phi$ to a landmark $\theta_{n_t}$ from wheelchair position $s_t$ are calculated in (13). More elaborate discussions of the EKF used in FastSLAM can be found in [7].

### C. Navigation

The idea of the navigation module is to find an optimal route between this location and a given destination in a global manner. With this we mean it needs to be able to plan a general route, but the navigation in itself is not responsible for avoiding obstacles on the way. It will calculate a set of ordered waypoints that, when connected, provide the optimal route with the provided information. These waypoints are usually connected with a straight line, which doesn't reflect reality since there will be obstacles along this route that need to be avoided. The navigation module does nothing to avoid these obstacles, it only provides a general bearing, the steering around obstacles is a responsibility for another module. The algorithm we choose for routing is A* search [8].

A* search is a form of breadth-first search which calculates its cost with a heuristic. Rather than only having the cost of the path so far $g$, there is also a cost which is an estimate of how much the cost will increase from that node towards the destination. This cost is called the heuristic $h$. An example of what a heuristic cost could be is the Euclidean distance from the current position to the goal position. This way, the algorithm tries to keep the path cost small, but also takes the distance that has yet to be traveled into account, and will make the tree search a lot faster in most cases. The formula for the total cost is then: $cost = g + h$. This cost will be used to decide which node to expand next.

## III. RESULTS

In a complete setup, the wheelchair is able to drive around in a closed environment. The established actuators can make it turn in place and drive straight ahead. The planning and localization modules are able to intelligently update new controls for the wheelchair. Their performance must yet be optimized.

When turning, the wheelchair will deviate 2.1 % of the angle given by the steering module. This means that $\alpha_2 = 0.021$. When driving straight ahead, the wheelchair will deviate 1 % of the distance commanded by the steering module. This means that $\alpha_1 = 0.01$. This data is collected by letting the wheelchair perform a controlled movement a meaningful number of times. A calibrated accelerometer, with a maximum error of 3, was used to measure the angles.

Our implementation of the VFH algorithm has only been tested through simulation, in player-stage. It was tested using a map of an office environment. The simulated wheelchair then got objectives from a user it had to attempt to reach. The implementation avoids objects and tries to find a path to a certain destination. However, local minima and maxima can put the wheelchair in a loop, which means that he will do the same things over and over without realizing that it's not getting any closer to it's goal. The test results showed that obstacles are avoided. In addition, the destination can be reached if the following conditions are met: There are no local minima/maxima close by, the destination is not situated in a corner, and all objects are stationary. How close to an object the destination is allowed to be further depends on how the parameters were chosen. Changing these parameters will influence performance on other areas, such as avoiding collisions, so is not always a good idea.

In the localization, data association happens by matching simulated landmarks to newly discovered landmarks. The only feature for this matching is location. Due to movement noise, a simulated landmark and an actual landmark will not be at exactly the same location. If the simulated and actual landmark are near to each other, they will be matched. In our current implementation this matching does not happen often enough for FastSLAM to decide on a most probable path followed by the wheelchair. This can be solved by increasing the maximum distance landmarks will be matched, however, this could also increase incorrect matching. Another solution which seems more promising, is to use additional features. Depending on the features chosen this could greatly improve our localization.

Extracting landmarks from the laser range scanner has a time complexity depending on the number of scales used, $O(n)$ with $n$ the number of scales. On 128 scales the algorithm takes an average of $14.11$ seconds, while on four scales the algorithm takes only $0.48$ seconds on average. The landmarks differ only $3.49$ cm and $2.82$ degrees on average.

In subsequent scans, up to $80\%$ of matching landmarks

 **ISBN: 978-1-61208-235-6** 

are identified, but the localization is prone to errors. Since no other data association but location is performed, false new landmarks are usually created and most particles have the same, unlikely probability.

## IV. CONCLUSION

We successfully build an electric wheelchair that is able to drive in a way a computer program asks it to. A computer formulates a command and transfers it to an Integrated Circuit, which translates the command to power on the motors and emits a PWM signal. This PWM signal is transferred to the motors through H-bridges.

The computer program is a system that was made by combining existing algorithms that are are responsible for the localization, the obstacle avoidance, and the navigation. These algorithms can successfully plan a route to a known destination, using the data that the wheelchair's sensors detect. This system combined with the working hardware, is the base for the project to construct an autonomous wheelchair and is a good first step towards achieving this goal. However, there is still a lot of optimization needed.

The localization is inaccurate and unstable, we discoverd that it is impossible to do FastSLAM with only a laser finder and data association using only location as feature. Since the wheelchair uses this location for it's planning and it's obstacle avoidance, this needs to be improved upon. We believe this can be done by extended data association in the landmark detection. The data association is now only dependent on the location of the landmark to the vehicle. Extracting more features from the landmark should improve the association.

Although obstacle avoidance is working, it's not colliding with objects, the wheelchair's main goal is to get to it's destination. However, due to local minima and maxima the robot sometimes locks in a loop which stops it from getting any closer to it's goal. A simple solution can be to also keep track of the wheelchair's path, and if an action is repeated more than once without a better result, then the wheelchair needs to try something different.

Finally, the hardware is able to receive commands and do a reasonable attempt to carry them out. Although it is not necessary that these attempts are very accurate, we need to have a better insight in the error on them. In addition, researching this will also give a better insight in how to increase control accuracy.

## REFERENCES


[1] T. Luettel and M. Himmelsbach, "Autonomous Ground Vehicles Concepts and a Path to the Future ," *IEEE Transactions on Robotics*, 2012.

[2] U. Franke, D. Gavrila, S. Görzig, F. Lindner, F. Paetzold, and C. Wöhler, "Autonomous driving goes downtown," *Intelligent Systems and their Applications, IEEE*, vol. 13, no. 6, pp. 40–48, 1998.

[3] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stanek, D. Stavens, A. Teichman, M. Werling, and S. Thrun, "Towards Fully Autonomous Driving: Systems and Algorithms," *IEEE Intelligent Vehicles Symposium*, no. IV, 2011.

[4] E. Guizzo, "How google's self-driving car works," October 2011, [retrieved: August, 2012]. [Online]. Available: http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works

[5] H. Yanco, "Integrating robotic research: a survey of robotic wheelchair development," in *AAAI Spring Symposium on Integrating Robotic Research*. Citeseer, 1998, pp. 136–141.

[6] J. Borenstein and Y. Koren, "The vector field histogram-fast obstacle avoidance for mobile robots," *Robotics and Automation, IEEE Transactions on*, vol. 7, no. 3, pp. 278–288, 1991.

[7] M. Montemerlo and S. Thrun, *FastSLAM A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics*, B. Siciliano, O. Khatib, and F. Groen, Eds. Berlin: Springer, 2007, vol. 27.

[8] S. Russel and P. Norvig, *Artificial Intelligence A Modern Approach*, 3rd ed. Upper Saddle River: Prentice Hall, 2010.

[9] J. Crowley, "World modeling and position estimation for a mobile robot using ultrasonic ranging," in *Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on*. IEEE, 1989, pp. 674–680.

[10] J. Borenstein and Y. Koren, "Obstacle avoidance with ultrasonic sensors," *Robotics and Automation, IEEE Journal of*, vol. 4, no. 2, pp. 213–218, 1988.

[11] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *The international journal of robotics research*, vol. 5, no. 1, pp. 90–98, 1986.

[12] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 3, pp. 229–241, Jun. 2001.

[13] SICK, "Operating Instructions Laser Measurement Systems of the LMS100 Product Family," Waldkirch, pp. 1–110, 2009.

[14] ——, "Laser Measurement Systems of the LMS100 Product Family," p. 8, 2009.

[15] R. Madhavan and H. Durrant-Whyte, "Natural landmark-based autonomous vehicle navigation," *Robotics and Autonomous Systems*, vol. 46, no. 2, pp. 79–95, Feb. 2004.

[16] F. Mokhtarian and A. Mackworth, "Scale-Based Description and Recognition of Planar Curves and Two-Dimensional Shapes," *IEEE transactions on pattern analysis and machine intelligence*, vol. 8, no. 1, pp. 34–43, Jan. 1986.

[17] H. Asada and M. Brady, "The Curvature Primal Sketch," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 1, pp. 2–14, 1984.


 ISBN: 978-1-61208-235-6 

[18] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: A factored solution to the simultaneous localization and mapping problem," in *Proceedings of the National conference on Artificial Intelligence*. AAAI Press, 2002, pp. 593–598.

[19] M. Montemerlo and S. Thrun, "Simultaneous Localization and Mapping with Unknown Data Association Using FastSLAM," in *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*. IEEE, 2003, pp. 1985–1991.

[20] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge: The MIT Press, 2006.

# Wireless Sensor Network Protocol for Smart Parking Application Experimental Study on the Arduino Platform

Luis Ostiz Urdiain, Carlos Pita Romero
*Computer Engineering and Networking Department*
*Universitat Ramon Llull, La Salle*
*Barcelona, Spain*
*Email: luis.ostiz@gmail.com, carlus.pita@gmail.com*

Jeroen Doggen, Tim Dams, Patrick Van Houtven
*Department of Applied Engineering and Technology*
*Artesis University College Antwerp*
*Antwerp, Belgium*
*Email: jeroen.doggen@artesis.be, tim.dams@artesis.be, patrick.vanhoutven@artesis.be*

***Abstract*—The paper presents an Arduino-based wireless sensor network to monitor parking lots using a non-standard low-power energy-balanced system. The event-driven routing protocol follows the hierarchical clustering philosophy. Energy is saved by minimising the number of transmissions needed to forward information to the base station. The smart sensor platform is build using the popular Arduino development platform, Sharp IR distance sensors and nRF24 low-power radio modules. Our practical results show that this platform is easy to use, but not the most appropriate platform to develop low-power wireless sensor network applications.**

***Keywords*-Wireless Sensor Network (WSN); Smart Parking Application; Clustering; Arduino; nRF24.**

## I. Introduction

With the rapid proliferation of vehicle availability and usage in recent years, finding a vacant car parking space is becoming more and more difficult, resulting in a number of practical conflicts (e.g., time, environmental [1]). Using Wireless Sensor Network (WSN) technology, we propose a low power solution to this parking problem.

A WSN consists of a large number of smart sensors which form a multi-hop network by radio communication in sensor fields. They measure and process information gathered from the sensing area and transmit it to the data base station. WSN can be used in many fields such as environment monitoring, intelligent transportation and smart homes [2]. In our system, we monitor a group of ten sensors, each one detecting periodically if there is a car parked a parking lot. This information is forwarded to a sink node. The sink node itself is connected to a management center, which is not yet a part of this research. The management center can help other drivers to discover areas with available parking lots.

The sensors nodes are based on the popular Arduino platform [**?**]. Arduino is an open-source electronics prototyping platform with flexible and easy-to-use hardware and software. We are developing a network protocol on top of nRF24L01 radio modules [3]. We have taken some features from other network WSN protocols, in particular clustering hierarchical protocols, such as Low Energy Adaptive Clustering Hierarchy (LEACH) [4] and Threshold sensitive Energy Efficient sensor Network (TEEN) [5].

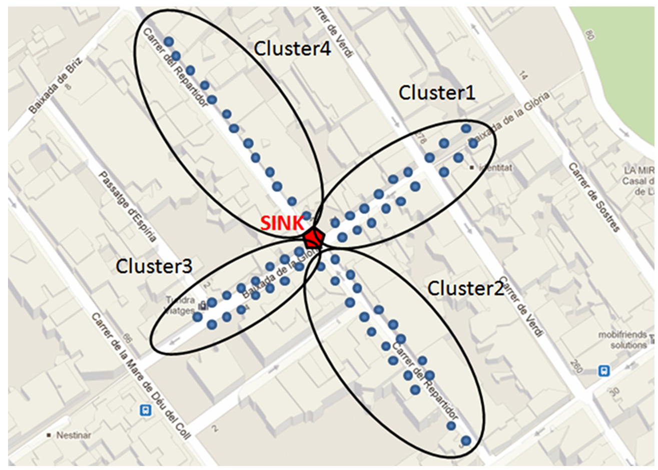



Figure 1. Proposed Single-Hop Clustering Topology

The remainder of the paper is organised as follows. In Section II, we discribe the system design the node hardware architecture and the software libraries. In Section III, we explain the design of the layer 2 and 3 WSN protocol operation. The research questions and experiments are in Section IV. Finally, we describe future work in Section V.

## II. System Design

For this experimental study we developed network of smart objects based on the Arduino development platform.

### *A. Cluster Network*

The proposed single-hop clustering topology is shown in Figure 1. The sensor network is divided in 5 groups, one for each cluster and one for the sink node. Each group uses a specific RF channel. The sensors detect if there is a car in the parking lot. The sensor will be located in a strategically fixed position on the floor and will sense periodically to detect cars. Knowing our nodes are deployed in a fixed position, we decided that the cluster members are invariable.

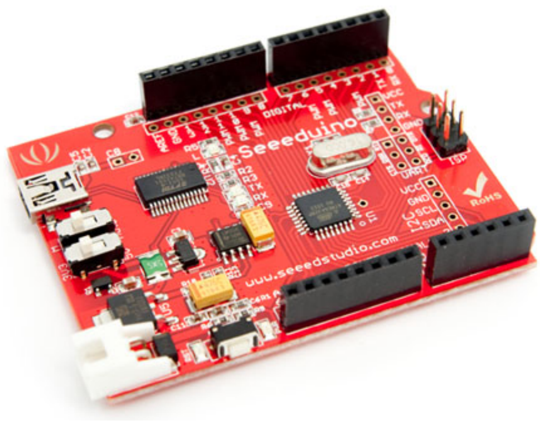



Figure 2. Seeeduino board

*1) Sensor Node (SN):* A SN periodically measures the distance with the sensor and compares it with the last measurement. If the result has changed it sends an advertisement message (ADV) to the cluster head.

*2) Cluster Head (CH):* The CH monitors a parking lot, but at the same time, it waits for advertisement messages from the other SNs. It aggregates the information and forwards it to the sink node. This role rotates between all the nodes within the same cluster.

*3) Sink Node:* This fixed node is continuously listening on designated channels to forward packets, coming from the cluster heads, to the management system.

### B. Node Architecture

We use the Seeeduino platform [6], an Arduino-compatible development board, but with many improvements on the hardware side, e.g., energy efficient surface mount device (SMD) components and extra analog and digital I/O pins. The Atmel ATmega328P microcontroller has 32 KB integrated flash memory with read-while-write capabilities, 1 KB EEPROM and 2 KB SRAM. We use a dedicated nRF24L01 low power radio module working in the 2.4 2.5 GHz ISM band. The radio uses FSK modulation type and can operate within 125 RF channels with 250 Kb/s, 1 or 2 Mb/s of data rate. We use the GP2Y0A21YK [7] infra-red proximity sensor made by Sharp, which is a wide-angle distance measuring sensor that delivers an analog output varying from 3.1 V at 10 cm to 0.4 V at 80 cm. We use 6 external 1.2 V rechargeable NiMH AA batteries (1300 mAh) providing 7.2 V per SN and a USB powered sink node. As the energy levels of the batteries decrease, the battery voltage goes down. In practice, the module stops working properly below the 6 V cut-off voltage.

### C. Software Libraries

The source code of all our libraries are publicly available under de GNU Lesser General Public License on their respective Google Code website [8].

*1) Low-Power Library:* This Arduino library, developed by Rocket Stream Electronics, allows us to put the microcontroller into different power saving modes [9].

*2) nRF24 Module Library:* This library for the RF module, developed by J. Coliz [10], provides the physical layer functions: transmitting and receiving packets. It also allows us to change the RF channel, set the retransmission options and power down the chip.

*3) Distance Sensor Library:* We developed this library to use Sharp IR Distance Sensors in Arduino projects. We use the sensor output to fetch the distance from a lookup table holding a sensor-specific transfer function.

*4) Cluster Network Library:* We developed this library that implements the layer 2 and 3 operation of the system. This library, described in detail in Section III is build on top of the nRF24 library.

*5) Detecting Car Library:* We assume that a car is occupying a parking lot when the distance value is between 10 cm and 60 cm for several seconds. If the distance value exceeds 60 cm, the parking lot is available. If the value is below 10 cm, the measurement is considered invalid.

*6) Node Energy Library:* Library which provides the current energy level depending on the time spent in the different phases. It is used to decide which node has the highest energy level during CH selection.

## III. WSN Protocol Design

We are developing and testing a non-standard wireless communication protocol based on some features of clustering hierarchical protocols such as LEACH [4] and TEEN [5].

Low Energy Adaptive Clustering Hierarchy Protocol (LEACH) is a very common WSN protocol. LEACH is based on the division of the network nodes in clusters and the election of a set of temporary CHs. The CHs are in charge of their part of the network and aggregate collected data for later delivery to a sink.

Our implementation is based on the LEACH clustering philosophy but instead of using a selection procedure based on probability, we use the remaining energy level to rotate the CH role periodically [11].

In TEEN, the environment is sensed constantly, but the transmission of the information to the CH is done only when the data exceed the upper and lower threshold previously defined by the CH. Therefore, this protocol reduces the number of the transmissions to the necessary, only when the measured parameter value is outside the range of interest.

Our measurement model to detect cars is based on this concept. A sensor only transmits the parking lot status when it detects an arriving or a leaving car. We can assume a sensor will not have new information to report very often. Thereby, we lower the number of necessary transmissions to a minimum, thus increasing the system's lifetime.

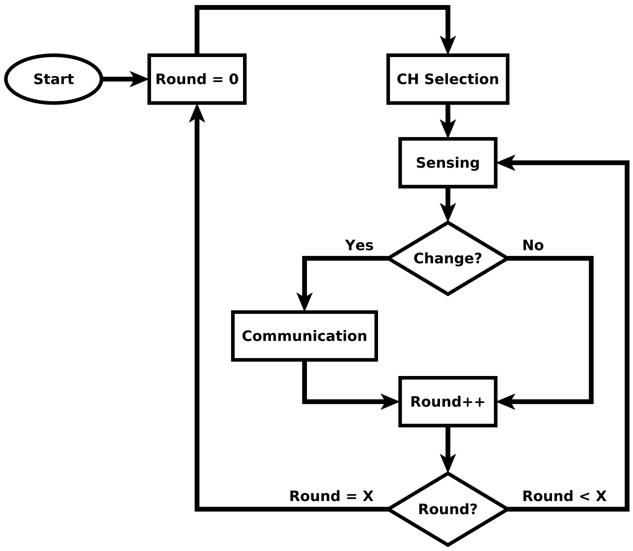



Figure 3. Flowchart of System Operation

### A. Layer 2 Operation

In the collision avoidance mechanism we implemented, every node that wants to transmit must first listen to the channel. If it is free, the node transmits the packet and waits for an acknowledgement. If it is occupied, the node waits for a back-off time and listens to the channel again. After a number of unsuccessful tries, the packet is dropped. The optimal amount of tries is still being evaluated.

### B. System Operation

The system operation, shown in Figure 3, is divided in three phases: CH selection, sensing, communication.

*1) Cluster Head Selection Phase:*

- The CH broadcasts an Energy Request message.
- The SNs measure their energy level and send it to the CH with an Energy Reply message.
- The CH collects all replies and compares the energy levels. The node with the highest energy level is selected as the new CH.
- CH broadcasts the new CH ID to all the SNs.
- The SNs update the CH ID at the same time.

*2) Sensing Phase:* The node detects if there is a change in the parking lot status. If there is one, it goes in to the *Communication Phase*, otherwise the node goes to sleep mode. Because the SNs are synchronised during the CH selection, they hould wake up at the same time.

*3) Communication Phase:* The node report the updated parking lot status to the CH by sending an ADV message. Meanwhile, the CH collects the ADV packets of all SNs. The CH sends an aggregated data packet (DATA), which contains all the collected status updates, to the sink.

## IV. EXPERIMENTAL STUDY

Our goal is to develop an energy-efficient implementation of a smart parking application. Some of the research questions in this project are:

1) Is the Arduino platform a good choice for this specific WSN application?
2) How can we minimise energy consumption per node to maximise network lifetime?
3) How do we optimise the CH selection algorithm?
4) How reliable is this system?

To evaluate the performance of the protocol, we propose two different variations of the same practical scenario. The sink is positioned in the middle of a square area and 10 SNs are positioned around the sink.

1) The whole network has been divided in 2 clusters, with 5 SNs per cluster.
2) The network is formed by 10 SNs which are part of the same cluster.

### A. Energy Consumption

To minimise the energy consumption, we maximise the sleep time by avoiding needless idle time and communication. To know how much energy is consumed per node, we defined different functional states depending on different energy hardware power requirements based on the low-power library specifications [9]. We measure the current consumption in each state to calculate the total energy consumption.

These defined states are: transmit, receive, idle and sleep.

First experiments on the energy consumption of the module showed a very inefficient sleep mode because the distance sensor was always powered up. An option to disable the sensor will be added to the distance sensor library.

### B. Packet Loss Ratio

A set of counters were added in order to record the number of packets transmitted successfully and packets lost. These measurements were made using a 90-minute recording time with a bit rate of 1 Mb/s. In the first case the packet loss was 41 %, in the second case 72 %. Further research is needed to solve this issue. A possible cause of this very high packet loss ratio is bad synchronisation.

### C. Synchronisation

The ATmega328P internal oscillator has a significant error margin when the board wakes up from sleeping mode. Because this error is different in each board, the SNs will not be able to communicate with the CH when a few rounds have passed. We implemented a non-optimal software-based solution around this hardware limitation:

1) The first round after a new CH is selected, the SNs wake up 4.8 seconds before the CH ($24 \times 0.2 = 4.8$).
2) Once the CH wakes up, it broadcasts a SYNC message.

3) The SNs receives the SYNC message from the CH calculated the time between waking up and the SYNC message.
4) The SN calculates its own sleeping time by taking the waiting time for the SYNC message in to account and adding one second as a safeguard.
5) In the second round after the CH selection, the procedure is similar, with a safeguard time of 100 ms.
6) From now on, the non-CH sleeps the calculated time.

## V. Conclusion and Future Work

We proposed a low-power WSN solution for a smart parking application. Along its development, the work has led to new challenges and opportunities that might be interesting to address.

### A. Further Development of the Application

The next step in the development of a usable application is server-side application that manages all sensor information, such as its geographical position, parking lot types and parking lot status. The user would be able to connect to the server using a website or with mobile phone application to request the location of available parking lots. Another useful idea is developing an application to facilitate monitoring of the smart parking network. Currently we are developing an embedded display module for in-car placement similar to [12]. This module sends requests over the Internet to the central server to find free parking spaces. When the car arrives in the correct street it can request up-to-date information from the CH.

### B. Protocol Enhancement

In order to polish the protocol, we propose to improve its scalability and enable more complex data aggregation. In a next phase it might be useful to have multi-hop communication and location aware CH selection [13]. One important aspect that is currently not considered is security. It could be a really interesting to introduce security techniques which would not affect the protocol's and the system's behaviour.

### C. Conclusion

We proposed a practical implementation of an event-driven WSN clustering protocol. Our protocol is hierarchical, all nodes are divided in clusters with one CH. We implemented an energy-aware CH selection algorithm similar to the LEACH protocol.

The system was implemented using Seeeduino development boards, nRF24L01 low-power RF modules and Sharp IR distance sensors. We implemented a software based synchronisation mechanism to solve the problems caused by the inaccuracy of the Arduino internal Timer.

These results show that the battery, sensor and radio are not the optimal hardware choices for this WSN Application. Although Arduino is easy to use as an experimental open-source platform, it is currently not the most appropriate platform to develop low-power WSN applications.

## References


[1] M. Idris, Y. Leng, E. Tamil, N. Noor, and Z. Razak, "Car park system: A review of smart parking system and its technology," *Information Technology Journal*, pp. 101–113, 2009.

[2] Y. Li, M. Thai, and W. Wu, *Wireless Sensor Networks and Applications*, ser. Signals and Communication Technology. Springer, 2008, [Retrieved: August, 2012]. [Online]. Available: http://books.google.be/books?id=x1MR5Ct-rp8C

[3] Olimex Ltd., "nRF24L01 Single Chip 2.4 GHz Transceiver Datasheet," [Retrieved: August, 2012]. [Online]. Available: http://www.olimex.com/dev/mod-nrf24L.html

[4] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocols for Wireless Microsensor Networks," in *33rd Hawaii International Conference on System Sciences*, 2000, pp. 1–10.

[5] A. Manjeshwar and D. P. Agrawal, "TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks," in *International Workshop on Parallel and Distributed Computing Issues in Wireless Sensor Networks and Mobile Computing, San Francisco, USA*, 2001, pp. 2009–2015.

[6] "Seeeduino v2.21," [Retrieved: August, 2012]. [Online]. Available: http://www.seeedstudio.com/wiki/Seeeduino_v2.21

[7] Sharp, "Sharp GP2Y0A21YK IR Distance Sensor Datasheet," 2005, [Retrieved: August, 2012]. [Online]. Available: http://www.sharpsma.com/webfm_send/1208

[8] J. Doggen, L. Ostiz, and C. Pita, "Software Libraries for the Arduino Platform," [Retrieved: August, 2012]. [Online]. Available: https://code.google.com/u/104098523773938750140

[9] Rocket Scream Electronics, "Lightweight Low Power Arduino Library," 2011, [Retrieved: August, 2012]. [Online]. Available: http://www.rocketscream.com/blog/2011/07/04/lightweight-low-power-arduino-library

[10] J. Coliz, "nRF24L01(+) 2.4 GHz Wireless Transceiver Driver," 2011, [Retrieved: August, 2012]. [Online]. Available: http://maniacbug.github.com/RF24/

[11] K. Ramesh and K. Somasundaram, "A comparative study of clusterhead selection algorithms in wireless sensor networks," *International Journal of Computer Science & Engineering Survey (IJCSES)*, pp. 153–164, 2012.

[12] V. P. Thakare and N. A. Chavan, "Performance evaluation of parking guidance and management system using wireless sensor network," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 1, pp. 96–102, 2012.

[13] T. V. Bhuvaneswari and V. Vaidehi, "Enhancement techniques incorporated in leach- a survey," *Indian Journal of Science and Technology*, vol. 2, pp. 36–44, 2009.

# Integrating Tiny Heterogenous and Autonomous Data Sources

Kim Tâm Huynh
Prism Laboratory,
University of Versailles,
Versailles, France
kim-tam.huynh@prism.uvsq.fr

Béatrice Finance
Prism Laboratory,
University of Versailles,
Versailles, France
beatrice.finance@prism.uvsq.fr

Mokrane Bouzeghoub
Prism Laboratory,
University of Versailles,
Versailles, France
mokrane.bouzeghoub@prism.uvsq.fr

***Abstract*—In this paper, we address the problem of integrating many heterogeneous and autonomous tiny data sources, available in an ambient environment. Our goal is to facilitate the development of context-aware and personalized embedded applications on mobile devices. The originality of the approach is the new ambient mediation architecture, which provides declarative and dynamic services, based on rules/triggers. These services provides facilities to develop and deploy ambient applications over devices such as smartphones. This paper reports on our first experiment prototype, combining Arduino+Android, in using such ambient mediator for an intelligent home application.An embedded mediation system CAIMAN is proposed and illustrated through a simple scenario.**

***Keywords*—*ambient data;embedded system;mediation system.***

## I. Introduction

Over the last 20 years, new paradigms such as ubiquitous computing, pervasive computing, ambient intelligence (AmI) have emerged with the development of wireless networks and the miniaturization of hardware components. Augusto and McCullagh [1] characterized AmI as “a digital environment that proactively, but sensibly, supports people in their daily lives”. The challenges posed by these paradigms are addressed by several research communities (networks, multi-agent systems, databases, Human-Machine Interface).

Today, we are witnessing an unprecedent explosion of mobile data volumes, i.e., ambient data. In 2011, 1.08 billion of mobile phone users have a smartphone. Smartphones as well as computers cannot really sense the world. In AmI environments, there is a need for tools for sensing and controlling more of the physical world . This is the role of the Arduino platform that can sense the environment by receiving input from a variety of sensors and can affect its surroundings by controlling lights, motors, and other actuators.

More and more ambient applications are developed for mobile environments, e.g., Waze [2], APILA [3]. Unfortunately, they are often developed from scratch, which is time and money consuming, and makes the software evolution quite difficult, in particular because components updates are frequent. The lack of a data management system for AmI does not ease the development of applications.

Ambient data have specific characteristics, they arrive as streams or as alert/notifications. Moreover, data are only relevant for a period of time and their interpretation depend on the user’s context and the user’s preferences. For instance, an information about a free parking place can be relevant for a user if this information is recent and if the parking place is nearby the user’s location. Another example is the heat setting to the right temperature in the room where a given person is and accordingly to his preferences. Such data streams are relatively small in their length/size.

For managing and integrating data streams, the database community has proposed two paradigms: DSMS (Data Stream Management System) processing [4] and Sensor Databases such as TinyDB [5]. DSMS is an evolution of the traditional DBMS (Data Base Management System). In DBMS, data are stored and users issue one-time queries on stored data. In others words, data are permanent and queries are transient. DBMS are not suitable for data streams. Indeed in DSMS, data are transient and queries permanent since they are continuously evaluated over the transient data, they are called continuous queries. A language CQL, has been proposed for managing and filtering data streams in a declarative way. Generally, these systems are centralized or clustered and assume that the data sources, i.e., sensors send continuously their data in a push mode, towards the DSMS. This assumption works if the data sources are known in advance, their schema does not change, they are always connected to the ethernet and do not have limited power. Indeed the push mode consumes a lot of energy, that’s why sensor databases have been proposed such as TinyDB. In this paradigm, we assume a network of sensors, data is acquired in a pull mode to avoid battery consumption. The query, i.e., Tiny SQL, is sent through the network and evaluated in a distributed mode. Sensors are active only when they have to answer a query. The advantages of this approach is that it is well adapted to the specificities of material and their constraints. The sensor network can contain a large number of sensors. However, the sensors are homogeneous, they all have a TinyOS and there is no mechanism of source discovery because the sensors are known. These two paradigms are not context-aware and cannot take dynamically into account heterogeneous and autonomous tiny data sources.

In this paper, we propose an ambient data mediation system which offers contextual and personalized data integration over autonomous and heterogenous tiny data sources such as

 ISBN: 978-1-61208-235-6 

smartphones and sensors/actuators, in a declarative way.

The plan of the paper is as follows. In Section 2, the motivations and the requirements for such a mediator are given. In Section 3, the ambient mediation approach is described. Finally, in Section 4, an application scenario is designed on top of our mediattor to illustrate our approach.

## II. Motivations and Requirements

For deploying ambient applications, there is a need of an ambient data mediation system (ADMS), which allows interoperability between a set of dynamic and loosely-coupled ambient data sources. An ambient data source is a (fixed or mobile) communicating object, which generates or consumes continuous (or discontinuous) flows of data. Among such objects, we can distinguish a wide spectrum of sensors and mobile phones. In addition to these data sources, there exist other ambient objects called actuators, that do not exchange data, but simply perform some actions, e.g., a led. Notice that a single physical object can play both the role of a data source and actuator. All ambient physical objects are abstracted by software services, which encapsulate them and make visible their capabilities, especially their data exchange protocol.

An ambient information system (AmIS) is a set of data flows provided by a collection of ambient objects to achieve the needs of AmI applications, e.g., intelligent home, health care. AmIS Objects may communicate between each other based on various communication protocols. For instance, sensors/actuators micro-controllers only offer a Wire-two-Wire Interface (TWI/I2C) for sharing data over a net of devices or sensors. On the other side, smartphones can exchange data in a more elaborate way. Some AmIS objects can play the role of a mediator, which is able to integrate and interpret data of many ambient data sources, as well as to perform actions over their environment. Most of the AmIS data may persist only a few seconds or minutes in the system, unless the application or the user decide otherwise for various reasons.

The main specific requirements imposed to the design of an ADMS are the following:

- Data sources are heterogeneous. They may be fixed or mobile and arbitrarily connected and disconnected from the mediator, during variable intervals of time. Data sources have different capacities in terms of storage and computation.
- The mediator can dynamically connect to the sources when and as long as they are active, i.e., visible over the wireless network and ready to provide data.
- The mediator should provide, for each application, the capability to define its data requirements in terms of event types, so offering similar concept as a mediator virtual schema, and a mechanism, which handles continuous queries.
- The mediator should be able to aggregate data flows originating from the same source and integrates data flows originating from different sources.
- The mediator should adapt itself to the user's context by continuously searching for the appropriate data sources, e.g., depending on the location and the time. It should also satisfy user's preferences in terms of data delivery, relevance to domain of interest, privacy.
- The mediator should be aware of energy consumption and manage consequently the connections to the sources and the usage of its resources.

These requirements clearly distinguish an ambient mediator from a conventional one [6].

## III. The Ambient Mediation Approach

We are currently designing and implementing an ADMS, called CAÏMAN for Context-aware dAta Integration and Management in Ambient eNvironments. The overview of the CAÏMAN architecture is depicted in Figure 1. Our aim is not to provide a complete set of data management services but rather a limited set of necessary functionalities to support the design of ambient applications that fit into lite clients such as smartphone, and exploit ambient data. The first goal is to provide a high-level declarative approach, based on ECA (Event-Condition-Action) rules/triggers [7], which permits user applications to interoperate over distributed ambient objects. The second goal is to facilitate object discovery and to handle dynamic connection/disconnection to these objects. The third goal is to make the ADMS aware of the user's context and user's preferences. For example, when a battery of a given equipment is low or when a user is too busy and does not want to be disturbed by anything, the rule processor should stop.

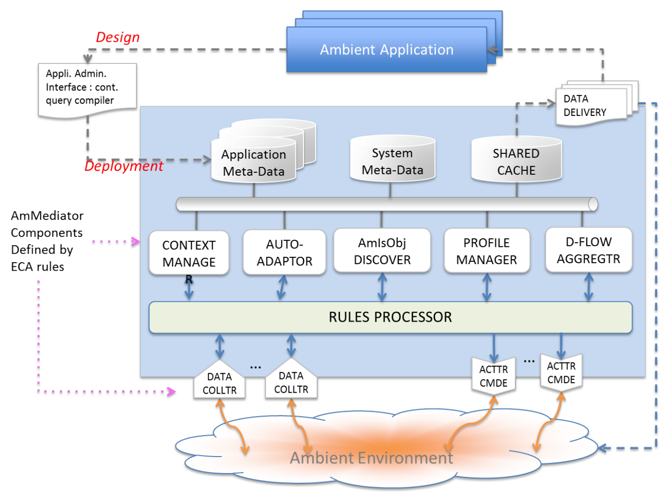



Fig. 1. CAÏMAN

The following subsections give an informal description of the main components of CAÏMAN.

### A. The Ambient Mediation Schema

The CAÏMAN mediation schema is defined as a set of events types, corresponding to the data flows required by ambient applications. An event type can be either simple *(SE)* or complex *(CE)*. A complex event type is a combination of other simple or complex events types.

Each ***event type*** *(SE & CE)* is defined by a set of attributes:

- *name*: name of the event type,

- *lifespan:* default time interval during which the event instance is valid,
- *aggrFunction*: function, which aggregates events to produce a complex event. For simple events, there is no aggregation function.

Each ***event instance** (SE & CE)* is defined by a set of attributes:

- *value*: event instance value,
- *source*: source name that captures the event instance,
- *raisingDate*: moment when the event instance is produced/observed by its source,
- *systemTime*: moment when the event instance is detected by the mediation system,
- *lifespan:* time interval during which the event instance is valid after its RaisingDate,
- *raisingLocation:* geo-location where an event instance is produced/observed by its source.

The *lifespan* is a metadata, which can be provided by the event source or assigned by the application. Event instances are relevant during a limited period of time. Pervasive environments can cause delays between the raising date of an event and the time of treating this event. The *raisingLocation* is a very useful notion for many location-aware applications. Indeed, the location can influence the relevance of a given event instance. For example, an event “flood” detected far away from a user can be irrelevant for him.

Once event types are defined, the application designer should specify how and when event instances are created or captured. This is done by specifying event detectors with windowing function. Depending on the event type and on the target data source, an event detector may be defined in various ways: a listener, a lookup function or any other procedure able to transform a specific signal into a semantic event. Finally, a set of continuous ECA rules is defined.

### *B. Binding ambient resources to the mediator*

In conventional mediators, data sources are known and linked once for all to the mediator at design time. In the context of an ambient mediation system, data sources are not known in advance but dynamically discovered at run time.

Ambient data sources are pervasive services, which may connect and disconnect arbitrarily, hence a centralized catalog of resources is useless. Only active objects in a given context are visible to the mediator. The *Resource Discovery* service is defined as a seeking function, which detects the surrounding active objects and establishes connections to them (called dynamic bindings). Binding a given data source to the mediator consists in matching the source meta data against a part of the mediation schema. If the match succeeds, it means that the data source can provide information to applications running over the ambient mediator, otherwise the remote source is considered as useless. A binding is then defined as a set of contextual mappings. The services provided by the mediator are then dependent on the successful mappings retrieved in the current location at a certain date. One of the main issues of the discovery process is to guarantee a continuous service even if data sources disconnect frequently. Besides the bindings, another issue that should be considered is data transformation. The data provided by a source is not necessarily compatible to the coding, format, unit and scale of the expected data at the mediator level. Data transformation is then another important functionality of the mediator. Source binding and data transformation services form what we call a data collector.

### *C. ECA Rules Processor*

Another fundamental service of CAÏMAN is the rule processor. Indeed, one of the main feature of our ADMS is its capability to provide a declarative language, which allows to describe most of the system semantics and the application semantics. This declarative language is the ECA rules language. User applications and mediation services are then defined by ECA rules. Each rule is defined using one or several event types defined in the mediation schema. The rule processor is an idempotent service to which ECA rules are submitted to be evaluated as long as event instances are produced by the application or the mediator. The rule processor has an operational semantics, which is clearly specified by various parameters such as event consumption and coupling modes.

### *D. Others components*

The *Application Metadata* contain event types, ECA rules, the context model, and the default user profile defined by the designer. These information are necessary for the different components. By using the context model, the *Context Manager* computes the current context, which can be used by the *Profile Manager* to infer the active profile, i.e., all user profile information, which are valid for this context.

Concerning the data, once the source is discovered and the data transformed by the *Data Collector*, the mediator proposes a *DataFlow Aggregator* component to process these flows of data and aggregate them. After executing the application rules, the *Data Delivery* component can deliver the result to the application or its ambient environment. The mediator can also execute actions through the *Actuator Command* component.

## IV. The Application Scenario

In this section, a specific scenario is chosen in order to illustrate and demonstrate our approach. Let us consider an ambient application scenario that wants to automatically control the air conditioning in the room where the person is, accordingly to his preferences. The user is mobile and can move from one room to another while keeping around the right air conditioning. For simplicity, we assume that the user is alone in the room. For doing so the application is constantly checking its environment to find if the room is well equipped, if it contains either a sensor of temperature or of humidity, and if an air conditioning actuator is present.

We first describe the ambient environment, which is composed of heterogeneous ambient data sources. Then, we illustrate the task that should done by our application scenario designer and what will happen when deploying it on top of our ambient mediation system.

### A. The ambient sources

In our ambient environment, we consider two sensors that capture respectively the humidity and the temperature and one actuator for the air conditioning. Each ambient data that is produced has a value and a timestamp, which corresponds to the time when the data has been captured. Each data source captures data at its own frequency. For each source corresponds a physical device characterized by an 'id', a 'type' and a version number. Each ambient source can export its capabilities in an XML document as depicted in Figure 2.

```
SOURCE 1 : (SENSOR)
<Metadata><Physical id=20 version=2.1  type= Arduino></Physical>
  <Sensor type= Humidity location=Room304 frequency=1000></Sensor></Metadata>
SOURCE 2 : (ACTUATOR)
<Metadata><Physical id=30 version=2.2  type= Arduino></Physical>
  <Actuator type=AirConditioning  location =Room304 >
   <action name= on ><parameter type=int name=duration></parameter></action>
   <action name=off ></action> </Actuator></Metadata
```

Fig. 2. Sources Description

### B. Design & Deployment

Most of the information that need to be specified by the designer are depicted in Figure 3 and explained in the following. First, the designer defines two simple event types: *UnvalidTemperature* and *UnvalidHumidity*. Both have a default lifespan of 2 min and no aggrFunction. Then, the complex event *UncomfortableSituation,* which is composed of the two simple events, is defined. The default lifespan is 5 min and an aggFunction Foo is associated. For each event, the designer must define a detector. In this scenario, simple event detectors are expressed declaratively in CQL-like manner and complex event detectors use a CEP-like language, composed of operators such as disjunction, sequence, etc. As said earlier, detectors can be defined in various ways.

The simple detector *DT* raises the *UnvalidTemperature* event when the temperature is not acceptable by the user. Due to space limitations, the *UnvalidHumidity* detector is omitted, since it is defined in a similar way. The complex detector *US* raises the *UncomfortableSituation* event when one of the simple event is raised within 50s. It uses the Foo aggrFunction for computing the event instance values. Notice that since data sources do not provide a lifespan, all detectors use the default value defined earlier by the designer.

For his application, the designer only needs the locality for his context model. For simplicity, we assume that the mediator already provides the detector function for this context. The contextual preferences of the user state that only sources and actuators that are located in the same room where he is, are accepted. So the designer provides the default profile for the application scenario by defining the resource discovery policies that are contextual. He also gives the domains of interest preferences such as the min and max temperature. Default values of the profile can be changed by the user at any time. Finally, the ECA application rule *MyScenario* is expressed, it consists in detecting an uncomfortable situation for the user and activating the air conditioning.

Once the event types, the detectors and the ECA rules are given, the application is compiled and deployed over the

```
Define SimpleDetector DT (EventType : 'UnvalidTemperature', Source : 'Temperature'){
        SELECT raise UnvalidTemperature( value : t.value, source : Source.name,
                        raisingDate : t.timestamp, systemTime : SYSDATE(),
                        raisingLocation: Source.location)
        FROM Temperature t [NOW]
        WHERE t.value not between (UserPref.Temperature.min, UserPref.Temperature.max);}

Define aggrFunction Foo (UnvalidTemperature t, UnvalidHumidity h) {
        raise UncomfortableSituation (
                value= (t.value or h.value); source = 'mediator' ;
                raisingDate = systemTime = SYSDATE(); raisingLocation = Context.Spatial.locality;)}

Define ComplexDetector US ( 'UncomfortableSituation', Foo('UnvalidTemperature', 'UnvalidHumidity') )
        { (UnvalidTemperature OR UnvalidHumidity) within 50s}

Profile
        Resource Discovery :
                source.location = Context.Spatial.locality; actuator.location = Context.Spatial.locality
        DomainsOfInterest :
                UserPref.Temperature : min = 16, max = 22; UserPref.Humidity : min = 20%, max = 40%
End Profile

Define ECA_Rule MyScenario
        EVENT : UncomfortableSituation[Now]
        ACTION : Activate ('airConditioning', 'on', '10 min')
```

Fig. 3. Design Phase

mediator. Once the application is started, the mediator creates the execution environment for the application. It activates the rule processor with the relevant ECA rules, as well as the complex event detectors. Once a relevant source is detected by the resource discovery component, a data collector is instantiated. It is responsible for the dynamic bindings. For instance, the right adaptor i.e., Arduino, is selected. All simple event detectors corresponding to the type of the data managed by the source are activated, e.g., temperature, humidity. The data collector requests data from the source. Then, event instance streams enter the mediator and are processed. When a source disappears the mediation removes the data collector instance, which in turn deletes all unnecessary flows and simple event detectors associated to the source.

## V. Conclusion and Future Work

In this paper, we have presented the requirements of AmI applications and proposed a mediation system CAÏMAN. Our system has been illustrated with a simple scenario. Its goal is to facilitate the development of embedded applications on mobile devices, that integrate ambient data which are different from conventional data. The approach is declarative. Its originality is to take into account, during the rules evaluation, the context and the user preferences. Application rules can be parameterized by a user so as his smartphone could be adapted to his personal needs. Ambient sources are fully implemented on Arduino boards and export their XML capabilities. The mediator is still under development on an Android smartphone. Performance evaluation still remains to be done.

## References


[1] J. C. Augusto and P. J. McCullagh, "Ambient intelligence: Concepts and applications," *ComSIS*, vol. 4, no. 1, pp. 1–27, 2007.

[2] (2012) The Waze website. [Online]. Available: http://www.waze.com/

[3] (2012) The Apila website. [Online]. Available: http://www.apila.fr/

[4] A. Arasu, B. Babcock, S. Babu, M. Datar, K. Ito, I. Nishizawa, J. Rosenstein, and J. Widom, "Stream: the stanford stream data manager (demo description)," in *Proc. SIGMOD'03*, 2003, p. 665.

[5] S. R. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "Tinydb: an acquisitional query processing system for sensor networks," *ACM Trans. Database Syst.*, vol. 30, no. 1, pp. 122–173, 2005.

[6] G. Wiederhold, "Mediators in the architecture of future information systems," *Computer*, vol. 25, no. 3, pp. 38–49, Mar. 1992.

[7] N. W. Paton and O. Díaz, "Active database systems," *ACM Comput. Surv.*, vol. 31, no. 1, pp. 63–103, 1999.

# User Privacy in Applications for Well-being and Well-working

Requirements and approaches for user controlled privacy

Wouter Bokhove, Bob Hulsebosch
Novay
Enschede, The Netherlands
wouter.bokhove@novay.nl, bob.hulsebosch@novay.nl

Bas van Schoonhoven, Maya Sappelli
TNO
Delft, The Netherlands
bas.vanschoonhoven@tno.nl, maya.sappelli@tno.nl

Kees Wouters
Philips Research
Eindhoven, The Netherlands
kees.c.b.a.wouters@philips.com

***Abstract*—Well-being applications at work and at home are expected to help people to continue contributing to society, the marketplace and the economy. To make them adaptive and intuitive, and allow them to provide personalized information and coaching to the user at the right time requires the availability of context information. The use of sensory devices for this purpose gives rise to an increased information level about users but also poses an increased privacy risk, especially when ubiquitous sensors and devices are networked and connected to on-line services. This paper describes user-centric approaches for protecting the privacy of users when applications use sensor data. Moreover, it assesses the compliance of these approaches with requirements for user controlled privacy and their suitability for well-being and well-working applications. Based on this assessment a number of privacy control approaches have been selected that are suitable for well-being and well-working applications.**

***Keywords - privacy; context; user-centric; control***

## I. Introduction

Well-being applications at work and at home assist individuals to continue their participation in society. In this paper, we advocate that user-centric sensing and reasoning techniques improve the efficiency and acceptability of applications for physical and mental well-being (mostly in a private context) and well-working (in a work context). These applications exploit information that describes the current context of the user. The availability of up-to-date contextual information enables application-developers to develop adaptive personalized apps.

However, this also results in privacy risks. The ubiquity of context information, and the relatively easy way of sharing this context information, increases the privacy risks for users. Data collection and data processing with respect for privacy [1] and data protection [2, 3], especially with regard to user awareness and control, are essential for privacy preservation and ultimately for the acceptance of well-being and well-working services. It is, therefore, important that the user remains in control of the collection, processing and distribution of data that is related to this user. In situations where user control is not feasible (e.g., in some professional or medical applications where context data is required to properly perform a job) users should still be informed properly about the way their personal information is handled.

These context-based reasoning systems make it virtually impossible for individuals to control access to privacy-sensitive information. Often, the user may not even know about the contextual information that has been sensed or inferred about her. In this paper, we argue that software frameworks that support context-management must obey two design considerations: *user control* and *usability*.

*User control* means that the user is empowered to decide what fragments of context information they consider sensitive, and in what situations they are prepared to share it with other parties for what purposes and under which conditions.

From the *usability* perspective, it is required that privacy control is user friendly and intuitive so users understand what they have specified, and are encouraged to actively use these controls instead of relying on default privacy settings.

Obviously, the trade-off triangle will play an important role in these considerations [4]: a balance should be found between ease of use, amount of control, and intrusiveness towards the user. In many cases, this results in conflicts.

Privacy control should be considered as a process of continuous adaption and conformation of preferences to the situational context and social practices. Key elements in this dynamic privacy control process are the ability for users to gain insight in and control over their current privacy settings and to get feedback on the impact of these settings given a situational context. These elements are fundamental to successful deployment of privacy-preserving well-being and well-working applications.

Improved privacy control and awareness might result in modifications of the privacy policies. These modifications can be considered as ad-hoc and executed manually by the user or automatically by the privacy control system and can be considered as part of a single loop learning cycle. As part of a so-called second or double loop, awareness could also result in a modification of the mindset of the user regarding her privacy policy settings, e.g., make them less tight.

 ISBN: 978-1-61208-235-6 

This paper discusses requirements for privacy control in context-aware services architectures. In addition, it presents the different functionalities needed to facilitate this control. The main objective of this control is to assist end-users in making decisions regarding privacy-sensitive information used by well-being and well-working applications. Although the actual privacy settings might be different for each application, the requirements for the controls are the same.

The structure of the paper is as follows. It starts with an overview of privacy control requirements followed by an inventory of known privacy control approaches is presented. Subsequently, each approach is assessed against the requirements. Finally, we summarize the outcome of this assessment and draw conclusions on the most suitable privacy control approaches for well-being and well-working applications.

## II. Privacy Control Requirements

Privacy architectures try to meet the fair information practices principles developed since the 1970s [5]. Since then, a lot of organizations have come up with privacy guidelines, directives, frameworks and/or principles to further specify or explain the privacy issues at hand and how these should be handled [1, 2, 3, 6, 7, 8, 9, 10, 11].

It has been recognized that implementation of privacy principles is especially difficult in ubiquitous systems involving (large) sensor systems which typically collect a lot of context information. Langheinrich [12] has tried to develop a comprehensive set of guidelines for designing privacy-aware ubiquitous systems based on a number of the aforementioned guidelines. Inspired by these privacy principles and focussing on user-centric privacy control, overview and usability, the following requirements can be distinguished.

### A. User-centric privacy control requirements

*1) Users must have privacy control over context information.*

Many users don't mind sharing personal information as long as they control how, where, when and with whom information is shared [13]. This is not only limited to static information like a user's name, birth date or more dynamic information like health records, status updates on social networks or the contents of emails, but is also applicable to a user's context. This results in the prime requirement that users must have privacy control over their context information.

*2) Users must be asked for permission at the time the context is requested.*

Users will not be able or willing to configure their privacy policies (completely) in advance. When context-information is requested, the user should be able to give or withhold an (informed) consent. Any solution for privacy control should thus allow for just-in-time (JIT) context requests. By allowing JIT consent requests, privacy policies must be applied in real-time.

This requirement is in direct contradiction with the usability requirement of unobtrusiveness (see also C1). A solution to reduce the invasiveness of JIT consent requests is to let the control system learn from responses and thus increasingly develop its privacy policies.

*3) Users must be able to modify and revoke their consents.*

When a user gives consent for accessing context-related information, the user should be able to revoke or modify this consent at any time. If consent is revocable, research [14] shows that this can reduce risk perception. In contrast to the current practice, where consents are mostly permanent (until revoked, if the user is able to find this option), it would be better to use access tokens with a limited life span or a limited number of uses.

*4) Users must have fine grained privacy control.*

The different attitude towards privacy can also be translated to a requirement with respect to the level of control that a user wants to exert with respect to privacy settings.

*5) Users must be able to define the granularity of the context information*.

Besides the level of detail of the privacy control settings, also the context information itself can be more or less detailed. Users should have control over this granularity. With respect to location, users may want to provide their exact GPS location or maybe just a (descriptive) derivative: home or work, neighborhood or city or region or country. Similar granularity levels (Quality of Context) can also be defined for other types of context information. This implies that the user is able to specify the granularity of the context information she is willing to share with a service provider.

### B. Overview requirements

*1) Users must be able to get an overview of all their privacy control settings.*

As users might be confronted with context-related privacy control issues throughout a long period of time, the user should have some way to get an overview of all their settings and consents, preferably in a single overview. Such an overview must provide insight into the users, applications and services that have access to a user's (aggregated) context information (and preferably also into the times and frequency this context information is accessed).

*2) Users must be able to get an overview of personal data provided to or accessed by a service*.

When users give consent to a service to access some of their personal data they may not be aware of the frequency this data will be used or the quality of context of the information. Therefore, it is required to have the possibility of an overview showing what data is used by which specific service. A step further would be to get insight in what is *derived* from the collected data by the context information consuming and/or aggregating parties.

### C. Usability requirements

*1) Users must not perceive privacy control as annoying or interruptive.*

Applications for well-being or well-working may need access to several different types of context information at different times and at different frequencies. The user should not need to grant or refuse access each and every time access is requested as this will make the control over context-related privacy settings a full-time job. Each time the user is asked to give permission this should be done in a manner which is neither interruptive nor annoying to the user. Therefore it should be done using a user friendly interface that enables an unobtrusive control of the privacy settings.

*2) Users must be able to understand the provided privacy controls*

Users must be able to understand what they give consent to, or put differently, the consent should be an informed consent. Informed consent is one of the requirements of the European Directive [2]. Since users have different levels of understanding and background knowledge, upholding this requirement is far from trivial.

*3) Privacy policies must be personalised.*

Many studies have investigated the attitude of users towards privacy issues. It is generally accepted to classify a person as being a privacy unconcerned (approximately 25%), pragmatist (approx. 50%) or fundamentalist (approx. 25%) [15]. Although this classification should not be used as a predictor for disclosing location information [16], it is found that users have a different attitude towards privacy.

## III. PRIVACY CONTROL APPROACHES

Multiple approaches to privacy control can be found both in literature and in the current practice of social networking sites. This section describes these approaches.

### A. Quality of Context

A form of obfuscation of context information is to alter its quality [17]. The assumption here is that detailed and specific context information is more privacy sensitive. From a privacy viewpoint, a user might want to restrict certain requesters from accessing very precise information.

### B. Symmetry

An important approach to maintain privacy in context-aware environments is the principle of minimal asymmetry, which in short states that the ability to obtain information should be coupled with the sharing of information between the data owner and consumer [18].

Balancing the amount of information flowing between peers is important to maintain the balance in any relationship. This is particularly the case for social relationships. Social systems often approach the symmetry principle by allowing the user to see the status of the other users she is connected with.

### C. Lying about yourself

Adapting data is a method for controlling what information that is sent out. A user can plan to lie or adapt the data after it is recorded and checked. This method reduces the tractability of the user’s actions. Another method for lying is by adding fake data to obfuscate the actual information.

### D. K-anonymisation / hiding in the crowd

In essence, the concept of k-anonymity relies on a simple protection mechanism: obfuscation. It then measures the provided privacy with a single parameter k. The value k determines the privacy protection in place: the larger the k is, the higher the privacy protection is. The k-anonymity scheme for location privacy has become popular, mainly due to its simplicity.

Another method for obfuscation of data is by hiding it in the crowd [19, 20]. This is a method based on k-anonymisation. By adding more or less random data (noise) to the signal it becomes more difficult to track down the original data. It can be seen as artificially creating other persons in the user’s region such that the conditions for k-anonymity are met automatically.

### E. Anonimysation and pseudonymisation

Pseudonymity is the ability to prove a consistent identity without revealing one's actual name, instead using an alias or pseudonym. Pseudonymity combines many of the advantages of both a known identity and anonymity. In anonymity, one's identity is unknown, but pseudonymity creates a separate, persistent "virtual" identity that cannot be linked to a specific person, group or organization. The purpose is to render the data record less identifying resulting in less customer or patient objections to its use. Data in this form is suitable for extensive analytics and processing.

Anonymity is often used as an underlying building block when implementing pseudonymity. In case of anonymity, no persistent name is used. It conceals the relationship between a particular user and the data about him. User model entries can no longer be assigned to a particular user, thus ensuring that they will remain secret. As a consequence, an anonymous communicating party cannot be remembered. It is also known as unlinkable anonymity.

### F. Consent

Consent is often required by legislation and is part of many fair information practices. By asking users for consent before sharing or accessing personally identifying information (PII) the user has great control over his privacy. In practice, however, most consents are based on ‘take-it-or-leave-it’ and thus leave little choice and control to the user with respect to his or her privacy. Several extensions to the ‘simple’ user consent questions can be defined, such as the option to decide which attributes will be released or the option to give only consent for a limited amount of time.

### G. Privacy control layers

When more control options become available it is likely to divide these options into layers where three layers are most common. Every layer contains more detailed settings. For controlling privacy settings, the top level is roughly suitable for users which are unconcerned about their privacy while the privacy fundamentalists can use the lowest level to configure their settings (almost) on policy level.

 ISBN: 978-1-61208-235-6 

### H. Fine-grained control

Research demonstrates that users have nuanced privacy preferences and that providing them with the ability to control personal information sharing based on more fine-grained and expressive privacy controls offers substantial benefit over simpler privacy controls.

There clearly is a need for greater expressiveness in privacy mechanisms, which control the conditions under which private information is shared on the Web [21]. Any increase in allowed expressiveness for privacy mechanisms leads to a strict improvement in their efficiency (i.e., the ability of individuals to share information without violating their privacy constraints), but comes at the cost of user friendliness as most privacy preferences will become relatively complex privacy.

### I. Grouping

Grouping attribute, people or service providers can help the user defining privacy policies.

#### 1) Grouping attributes

The clustering of several personal data attributes for which the same privacy policy will hold is a common way for current online services to organize consent of users. Clustering of attributes offers users a clear overview of which attributes will be shared and it provides a fast and easy way to give consent. However, in many current services users lack the possibility to cluster attributes themselves, or to alter the predefined clustering.

#### 2) Grouping people

Another way of clustering in privacy settings is to cluster people that have access to a particular attribute or several attributes. In the EU project PrimeLife [14], the social network Clique was developed which was based on this idea of clustering. Clustering makes the audience for users who see their information more transparent and it allows users to keep different parts of their identity separate (for example professional and personal life). This type of clustering thus allows for audience segregation [22]. Currently several social network sites such as Google+ (which named their groups ‘Circles’), Facebook (at which you can define multiple Lists of friends) have incorporated this clustering into their privacy settings.

#### 3) Grouping service providers

The third way in which privacy settings may be clustered is by arranging service providers in groups that may receive certain data based on certain characteristics of the service provider. A potential issue with this approach is the question who is determining in what cluster a service provider fits in.

### J. Removing Policies

Another privacy control option is to remove existing policies. Policies can be rules that the system has learned regarding consents the user have given. Kill switches exist that revoke all privacy settings at once and can be considered to be a batch version of the possibility to remove policies.

### K. Overview

Awareness starts with having an overview that captures the kind of information that is being shared with consuming parties (other users or service providers) under what conditions. At the moment this sharing information is far too scattered. Typically, consent is given once during installation of the application and forgotten afterwards. Having an overview of all consents given in the past to service providers that control certain personal data attributes would be an ideal starting point for privacy control. The size and complexity of the overview will strongly depend on the user’s privacy attitude: unconcerned, pragmatic, or concerned [15]. Overviews could include all given consents, which information is available to specific others or when or how often a service retrieves specific context information and exactly which information is retrieved. The overview may lead to an increased user awareness concerning her privacy settings and prevention of inadvertent invasions of privacy.

### L. Privacy Mirror

Privacy mirror is a method that makes the user aware of what information she is sharing and with whom she is sharing it with. It is a method for checking whether your privacy controls are working the way you expect them to work.

### M. Privacy Quiz

A privacy quiz can be used to make the user aware of her privacy settings. It ensures that the user understands what happened to his or her data. The privacy quiz can be implemented by asking the user to answer a privacy-related question. Depending on the complexity of the privacy policies, these questions can be very simple or more advanced. Answering the questions should be optional. With many policies there are a lot of questions possible that can automatically be generated. This is particularly the case if context is taken into account in the policy rules. When a user answers several questions wrong, she is expected to update here privacy settings.

### N. Notifications

Notifications are part of many privacy regulations and fair information policies and play an important role in raising and maintaining awareness with the user with respect to his or her privacy. The user can be informed of his personal data being accessed and used by a service provider in many different ways. For example, the user can be notified of each of the times a service provider accesses a certain piece of privacy information. In some cases, this will probably lead to the undesired situation in which the user is constantly being notified which will reduce the power of notifications in itself and the user intrusiveness is too large. The number of notifications can be reduced by notifying a user only when a service provider is accessing information in an unusual frequency or after a fixed number of times. It could also be envisioned that a user is notified only when a service provider becomes active after being dormant for some time. The opposite is also possible: a user might be notified when a service provider has been granted access to personal information, but has not actually accessed this information for some time.

*O. Making suggestions*

The information gathered from previous behavior regarding sharing of information of the user and choices of the user made regarding consent can be used to suggest privacy settings and specific privacy rules for future situations. These suggestions could include a number of previously mentioned privacy controls, such as clustering or time-based consent.

## IV. DISCUSSION

The many different approaches that were described earlier may all help in one way or another in increasing the level of awareness of, or control of users over, their personal information. However, it is impossible to simply implement all of these approaches, as this would result in inconsistencies, and may not be necessary to ensure adequate privacy protection to start with. To determine which approaches are preferred, they are first matched to the requirements and then their application for well-being and well-working applications is discussed.

Discussing how the individual approaches relate to each of the requirements is not feasible considering the large number of combinations that would need to be analyzed and discussed. Therefore, as a starting point of the analysis we mapped the requirements to the approaches. The result is shown in Table I. A '+' indicates that the approach can be used to implement a requirement, a '–' means that it conflicts with a requirement, and a '0' means that there is a dependency on the actual implementation. Empty cells indicate there is no relation between the approach and the requirement.

TABLE I. MAPPING OF PRIVACY CONTROL REQUIREMENTS TO THE DESCRIBED APPROACHES

| **Approaches** | **Privacy Control Requirements** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ***A1*** | ***A2*** | ***A3*** | ***A4*** | ***A5*** | ***B1*** | ***B2*** | ***C1*** | ***C2*** | ***C3*** |
| A. Quality of Context | + | + | | + | + | | | 0 | + | + |
| B. Symmetry | + | | | | | | | 0 | + | + |
| C. Lying about yourself | + | + | | | – | | | – | + | 0 |
| D. K-anonymisation | + | | | | + | | | + | – | + |
| E. Anonymisation and pseudonymisation | + | | | | | | | 0 | 0 | 0 |
| F. Consent | + | + | + | 0 | | | + | 0 | + | |
| G. Privacy control layers | + | – | + | - | | | + | + | 0 | + |
| H. Fine-grained control | + | – | + | + | | | + | – | – | 0 |
| I. Grouping | + | – | | + | | 0 | | + | + | + |
| J. Removing Policies | + | – | + | | | | + | 0 | + | |
| K. Overview | | – | | | | + | | + | + | |
| L. Privacy Mirror | | – | | | | + | | + | + | + |
| M. Privacy Quiz | | – | | | | + | | – | + | |
| N. Notifications | | + | | | | + | | – | + | + |
| O. Making suggestions | + | | | | | + | | – | + | + |

Based on the mapping presented in this table, several observations can be made. A number of approaches appear especially suitable for giving the user control. Asking consent to the user before sharing or processing his personal data implies that the user has control, and can control the handling of his personal data so that it matches his personal preferences. Most approaches that give the user control also allow a certain level of personalized control. Good examples that also provide the user with fine-grained control and fill-in a number of other requirements are: Quality of Context, Privacy Control Layers, or grouping of attributes, people, and service providers. More direct forms of control can for example be implemented using a Kill switch.

Providing the user with a good overview is essential, because consent can only be meaningfully given when it is informed consent.

User friendliness is an essential issue here, as some awareness approaches do not fit this requirement. This leaves approaches such as giving an Overview, or using a Privacy Mirror or Privacy Quiz.

The well-being and well-working applications aim at increasing physical and mental well-being of users. The unobtrusive nature of these applications that consume detailed privacy sensitive information to enhance service experience and effectiveness motivates the need for intuitive approaches able to cope with the high dynamic nature of situational changes. The approaches are therefore divided into three categories depending on how they are suitable for this application domain.

### A. Preferred approaches

Several approaches are very suitable for this application domain.

Quality of context is a very relevant tool, as information needs only to be as precise as required by an application, and no more. For example, if a person's heart is being monitored using ECG (ElectroCardioGram), it may be not be necessary to transmit and process the detailed (and more revealing) ECG data. Instead, a derived current heart rate may be sufficient.

Consent is legally and ethically a strong requirement and an essential precondition for the user to be in control. The way in which the user can give consent is important, however. For this, the use of other (more specific) privacy approaches is necessary.

Grouping attributes, people and service providers is especially relevant for well-being and well-working applications. Using (configurable) clusters of attributes, and service providers with which to share personal information, provides a level of control that may offer a suitable implementation of the "control" approach. A balance will have to be found, however, in the level of detail in which grouping takes place. Also, grouping has to be done before the actual processing is done, putting some limits on its use.

Privacy control layers are a way to allow users with different privacy attitudes to translate their personal privacy concerns in a convenient way. As users of well-being and well-working applications will have a diverse attitude and including context in privacy preferences may lead to complex, fine-grained control requirements, privacy control layers will be needed.

Overview is important for users to get an awareness on how their personal information is being processed. Moreover, for consent to be meaningful it needs to be informed consent, so the user must have an understanding of what information is shared with whom.

Notifications may be used to maintain awareness of what is happening with the user's personal information and is a suitable tool for use in well-being and well-working applications. It may also provide a way to give the user just-in-time control. Of course care must be taken not to "spam" the user with notifications that are not relevant.

### B. Conditional approaches

Some approaches are suitable only in specific situations or when specific conditions have been met.

Symmetry is a principle which is mainly relevant in sharing information with one's peers. So this may only be useful for some specific well-being and well-working applications, even though it is largely compatible with the requirements.

Anonymisation and pseudonymisation are powerful tools but may be difficult to successfully implement in some applications because of the kind of data that may be monitored. Essential data items in a well-being and well-working application include many potentially identifying features, such as age, gender, weight, health status, etcetera.

Fine-grained control can be useful for well-being and well-working applications as the number of options, especially when a large number of context sources is being used, will be quite high. Fine-grained control allows for the users of these applications to accurately control their privacy. Fine-grained control requires Privacy Control Layers for usability reasons.

Removing policies and a kill switch may be useful for providing the user with a high level of control in well-being and well-working applications, but depends largely on how this is implemented. As many well-being monitoring applications may depend on the long-term availability of data to discover trends, these may not be suitable control tools for some applications.

A privacy mirror is a potentially very usable tool for increasing user awareness, but this depends strongly on how it is implemented. Also, some personal information related to psychological or mental well-being may not be in a format that gives much insight in what is actually being shared (e.g. detailed sensor information).

Making suggestions is an approach that can be used to support users in making decisions on their privacy "policies". Although this fits with the well-being and well-working applications that aim at supporting the user in similar ways for other ends, it requires a large amount of privacy control settings (particular consents) before becoming useful. Making suggestions can be then used to optimize and to make privacy control more user-friendly.

Ask, but don't tell can be useful for users whose well-being is constantly being monitored as this approach allows them to get off the grid temporarily and thus allows the user to protect his or her privacy in a simple way. As medical or life-style advice is based on the observed information applying this tool frequently could have an adverse effect.

### C. Unsuitable approaches

Some approaches are generally not useful or difficult to implement in the well-being and well-working domain. We discuss these below.

Lying about yourself as a privacy approach in a well-being of well-working application environment may have very undesirable effects, as the quality of guidance provided by well-being and well-working applications depend on accurate information. Acting after medical or life-style advice based on incorrect information may have detrimental effects.

k-anonymisation depends on hiding the user's personal information in a large number of other user's personal information. However, as the well-being and well-working applications depend on providing specific users with feedback based on their specific personal information, this is not a useful technique for such applications. This is also true for hiding in the crowd, for the same reason.

Confronting the user with a privacy quiz is intrusive, and therefore useful only in specific circumstances, for example for privacy settings that are very important.

## V. CONCLUSIONS AND FUTURE WORK

This paper presented a number of privacy control approaches which are mapped to privacy control requirements. Also, the suitability for well-being and well-working applications is discussed. Based on this analysis it has become clear that controlling one's privacy with respect to context information, while finding the proper balance between being easy to understand for the end-user, being fine-grained and being unobtrusive, is not an easy task.

The logical next step would be to find out if it is possible to use the context of the user to automatically make decisions about sharing his or her context information, i.e. determine the 'context-awarability' of the various privacy control approaches. If this can be done, this will result in more user-friendly and adaptive solutions (e.g., the user will not be asked for consent while in an important meeting or while sleeping). Context-aware adaptive privacy might exploit the ability to sense and use contextual information to augment or replace traditional user privacy control mechanisms by making them more flexible, intuitive and less intrusive. Moreover, we also intend to determine which sensors will be best suitable for this purpose thereby taking into account the quality of the provided sensor information, reliability and its dynamicity. This is currently work in progress and will lead to the development and user-evaluation of well-being and well-working applications that takes into account several privacy control approaches that are context aware.

## ACKNOWLEDGMENT

This publication was supported by the Dutch national program COMMIT (project P7 SWELL). We thank the SWELL members Linda Kool, Saskia Koldijk, Maarten Wegdam, and Wessel Kraaij for the inspirational discussions.

## REFERENCES

[1] "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data", 2002.

[2] "Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data", EU, 1995.

[3] "Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector", EU, 2002.

[4] M. Wegdam and D.-J. Plas, "Empowering users to control their privacy in context-aware systems through interactive consent", Technical Report TR-CTIT-08-66, Centre for Telematics and Information Technology University of Twente, Enschede, 2008.

[5] R. Gellman, Fair Information Practices: a Basic History, April 25, 2012.

[6] Federal Trade Commision, "Privacy Online: Fair Information Practices in the Electronic Marketplace, a report to Congress", 2000.

[7] K. Cameron'Identity Weblog (www.identityblog.com), "The Laws of Identity", 2005.

[8] NIST, "Security and Privacy Controls for Federal Information Systems and Organizations", NIST Special Publication 800-53, Draft Appendix J, 2011.

[9] AICPA/CICA, "Generally Accepted Privacy Principles (GAAP)", 2009.

[10] "APEC Privacy Framework", 2004.

[11] U.S. Department of Commerce, "Safe Harbor Privacy Principles", 2000.

[12] H. Langheinrich, "Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems", Proceedings of the 3rd international conference on Ubiquitous Computing, pp.273-291, September 30-October 2, 2001.

[13] B. Schneier, "Google and Facebook's Privacy Illusion", 2010.

[14] EU, PrimeLife project, see: http://www.primelife.eu

[15] P. Kumaraguru and L.F. Cranor., "Privacy Indexes: A Survey of Westin's Studies", Institute for Software Research International, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, TR CMU-ISRI-5-138, 2005.

[16] S. Consolvo, I. Smith, T. Mathews and A. LaMarca, "Location Disclosure to Social Relations: Why, When, & What People Want to Share", in CHI '05: Proceedings of the SIGCHI Conference on Human factors in Computing Systems, New York, NY, 2005.

[17] K. Sheikh, M. Wegdam and M. van Sinderen, "Quality-of-Context and its use for Protecting Privacy in Context Aware Systems", Journal of Software, Volume 3:83–93, 2008.

[18] A. Kofod-Petersen, E. Klæboe, J. Jervidalo, K. Aaltvedt, M. Romnes and T.M. Nyhus, "Implementing privacy as symmetry in location-aware systems", Proceedings of the International Workshop on Combining Context with Trust, Privacy and Security (CAT 2008), volume 371, pp. 1–10, Trondheim, Norway, 2008.

[19] F. Bustamente and L. Amaral, "Software Improves P2P Privacy By Hiding In The Crowd", 2009.

[20] F. Li, J. Sun, S. Papadimitriou, G.A. Mihaila and I. Stanoi, "Hiding in the Crowd: Privacy Preservation on Evolving Streams through Correlation Tracking", ICDE, 2007.

[21] M. Benisch, P.G. Kelley, N. Sadeh, T. Sandholm, L.F. Cranor, P.H. Drielsma and J. Tsai "The Impact of Expressiveness on the Effectiveness of Privacy Mechanisms for Location Sharing", CMU-ISR-08-141, 2008.

[22] E. Goffman, "The Presentation of Self in Everyday Life", New York: Anchor, Doubleday, 1959.

# Development of the Future Body-Finger

A novel travel aid for the blind

Kiyohide Ito
Department of Media Architecture
Future University Hakodate
Hakodate, Japan
itokiyo@fun.ac.jp

Yoshiharu Fujimoto
Department of Media Architecture
Future University Hakodate
Hakodate, Japan
fujimonia@gmail.com

Junichi Akita
School of Electrical and Computer Engineering
Kanazawa University
Kanazawa, Japan
akita@is.t.kanazawa-u.ac.jp

Ryoko Otsuki
Department of Media Architecture
Future University Hakodate
Hakodate, Japan
g2112008@fun.ac.jp

Akihiro Masatani
Graduate School of Electrical and Computer Engineering
Kanazawa University
Kanazawa, Japan
masatani@ifdl.jp

Takanori Komatsu
Faculty of Textile Science and Technology
Shinshu University
Ueda, Japan
tkomat@shinshu-u.ac.jp

Makoto Okamoto
Department of Media Architecture
Future University Hakodate
Hakodate, Japan
maq@fun.ac.jp

Tetsuo Ono
Graduate School of Information Science and Technology
Hokkaido University
Sapporo, Japan
tono@complex.ist.hokudai.ac.jp

***Abstract*—We have developed a device with a haptic interface that enables blind persons to perceive the environment. The device, which we call the Future Body-Finger (FB-F), measures the distance between a person and an object using an infrared sensor and transmits distance information to the user's haptic sense by the angular motions of a link. In this report on the FB-F, we (1) outline its concept, (2) describe its mechanism, and (3) demonstrate its usability by comparing it to commercially available products.**

***Keywords-haptic interface; electric travel aid; perception***

## I. INTRODUCTION

Human beings are often said to obtain almost all of their information via the visual modality. However, nonvisual modalities can also provide rich information. Therefore, shedding light on the positive use of nonvisual modalities for exploring our surroundings is an important issue. In addition, we assume that a novel subjective experience of the environment (e.g., that of an “extended mind” [1]) could be realized by means of devices that function as parts of our body and endow us with “embodied cognition.” We have been working to develop devices that harness this form of cognition. We refer to a device that enables us to have an “extended mind” experience as a “Future Body.” As a first step, we have developed a device called the “Future Body-Finger” (hereafter, “FB-F”) that enables visually impaired people (hereafter, “the blind”) to perceive their surroundings.

Numerous devices called electric travel aids (ETAs) [2] have been introduced to assist locomotion of the blind. In order to ensure the safety of locomotion, ETAs have incorporated functions that obtain information on orientation. For example, ETA sensors determine a user’s location, the direction in which the user moves, and the distance of nearby objects.

ETAs are categorized into two types based on their output modality: auditory and haptic. Devices of the former type transform spatial information into audible sound [3]. For example, Tri-sensor [4] and Miniguide [5] measure distance via ultrasonic waves, convert the data into sound, and convey information on the distances of objects around a user. These devices emit a low-pitched sound when an object

 ISBN: 978-1-61208-235-6 

is distant from the user and a higher-pitched sound when the object approaches.

Devices of the latter type, which use haptic output [6], typically convert spatial information into vibration. The intensity and frequency of the vibration are conveyed as information via the user's skin sense. For example, the vibration frequency of such a device increases when a user approaches an object. Users become accommodated to the skin stimulation provided by this type of haptic output because the threshold of vibration sensitivity becomes increasingly high as users are exposed to the same vibration stimulus for increasingly longer times.

Unfortunately, the aforementioned ready-made devices are difficult to handle skillfully. For example, the blind need to develop higher cognitive abilities to comprehend the information conveyed by sound or vibration, i.e., they must interpret the sound and vibration signals by learning to associate a sound of a certain pitch or arbitrary vibration frequency with the distance to an object. Otherwise, they will fail to infer the locations of objects. In this sense, blind children will have difficulty managing these devices because their abilities in higher-order cognitive processing are not fully developed. Furthermore, adventitiously blind persons may have to make greater efforts to use ETAs because they have difficulty in discriminating sound pitch as well as the intensity and frequency of vibration compared with congenitally blind persons.

## II. CONCEPT OF FB-F: A "SMART" MECHANISM

As suggested by Runeson [7], we define a "smart" mechanism as a mechanism that directly registers complex variables. We adopted this concept to develop the FB-F. To provide an example of a "smart" mechanism, we begin by explaining how a polar planimeter works (Figure 1). A polar planimeter is a tool to measure the area of irregular shapes, which requires the calculation of complex variables. When a user moves the tracer arm, an attached measuring wheel carefully traces the outline with an index to calculate the area (a complex variable) automatically. The length and angle measured by the polar planimeter are directly proportional to the area. The device is so simple to use that those who have no knowledge of the calculations (e.g., summing up small pieces of a figure) can easily determine the area. Therefore, the advantage of this "smart" mechanism is that it does not require any calculating ability.

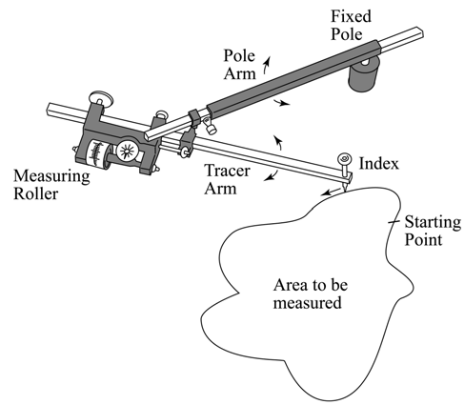



Figure 1. Polar planimeter

## III. HARDWARE CONFIGURATION

We designed the FB-F in order to solve the problems of usability mentioned in the previous section. FB-F is expected to enable users to obtain spatial information such as the direction and distance to an object intuitively according to the user's exploratory actions without the use of cognitive inference.

The FB-F is characterized by its haptic "smart" user interface. Users do not need mathematical calculations, mental inference, or higher-order cognitive processing to recognize the distance to an object, whether moving or not.

The hardware architecture of the prototype FB-F is shown in Figure 2. The developed FB-F consists of three functional blocks—a controller, sensor, and actuator units—which are connected to a common communication channel. Each unit has microcontroller (MCU, Cypress CY8C21123). The sensor unit has a position-sensitive device (PSD)-type distance sensor that radiates infrared rays toward an object; it detects the reflection position of the received rays by using a PSD that implements a trigonometric distance measurement technique. The microcontroller on the sensor unit calculates the distance from the FB-F to the object. The actuator unit has a servo motor equipped with a 55-mm-long lever to form a 1-DOF (one degree-of-freedom) link. The microcontroller on the actuator unit controls the servo motor according to the received angular information. The controller unit periodically requests distance information from the sensor unit, converts the measured distance to angular information, and transmits it to the actuator unit; this chain of operations forms the sensor-actuator system. The angle of the link increases when the distance between the FB-F and an object decreases (e.g., when an object approaches), whereas it decreases when the distance increases. Users hold the FB-F (Figure 3) and place their forefinger on the link. The finger bends or extends depending on the link's angular motion. The angle changes from 0 to 70° in correspondence with the metric distances between a user and an object. The FB-F can provide users with distance information because the extent that the user bends his/her forefinger is directly associated with the link movements.

 ISBN: 978-1-61208-235-6

The hardware specifications of the prototype FB-F are as follows: weight, 60 g; height, 7.5 cm; width, 4.5 cm; and depth, 3.5 cm. The distance measured varies from 1,000 to 2,800 mm as the link angle changes from 70 to 0°. The distance-angle coefficient is –7°/180 mm.

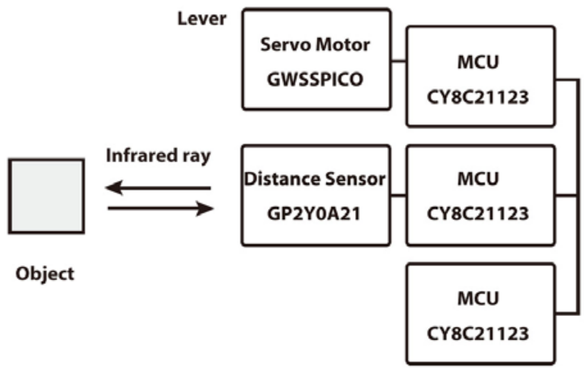



Figure 2. Block diagram of Future Body-Finger

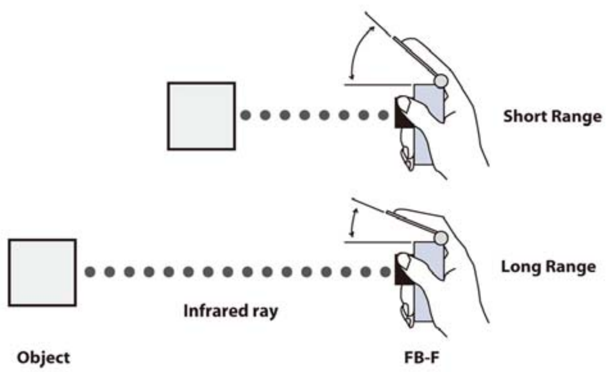



Figure 3. Operation of Future Body-Finger

## IV. PSYCHOLOGICAL STUDY ON ACCURACY OF ESTIMATED DISTANCE BY FUTURE BODY-FINGER

### A. *Purpose*

We conducted a psychological experiment to demonstrate that the FB-F enables users to perceive distance to an object more accurately than commercially available products.

### B. *Method*

Participants: 16 persons, blind or sighted, participated in the experiment. Eight blind adults, four congenitally and four adventitiously, participated in the blind group. Their ages were between 28 and 57 (mean = 43; SD = 9.50). Eight sighted adults (range, 20–22; mean = 20.75; SD = 0.89) participated in the sighted group.

Stimuli: A piece of cardboard adhered to a whiteboard (1.6 m × 1.0 m × 0.02 m) was used as the object. We used a standard stimulus and 5 test stimuli for the trials. The object was presented at a distance of 1.0 m from an FB-F device fixed on the table (“standard stimulus”) or at one of 5 positions ranging between 1.0 and 2.6 m from the table (“test stimuli”). The distance between each pair of adjacent test stimuli was 0.4 m.

Device Conditions: The type of ETA (the FB-F, a Vibratory device, and a Sonar device) was varied as a within-subject factor. Participants were asked to estimate the distance to stimuli under different device conditions (i.e., three types of ETAs). The Vibratory and Sonar devices were commercial products. The Vibratory device (7.0 cm × 4.0 cm × 2.5 cm) had a haptic interface that transformed measured distances into vibratory signals. The Sonar device (6 cm × 3.5 cm × 1.5 cm) transformed measured distances into audible sounds (i.e., sounds with a specific pitch). Both devices used ultrasonic waves to detect the distance to an object.

Procedure: Figure 4 shows the experimental setup. In each trial, participants were asked to use an ETA device to detect the distance to a stimulus that was presented for 3 s. Initially, the standard stimulus was presented, after which one of the five test stimuli was randomly presented. The magnitude estimation method was used to estimate the distance to the presented stimulus. In this method, a participant was asked to report the magnitude of a stimulus that corresponded to some proportion of the standard. The participant estimated his/her subjective experience by assigning numbers to the stimuli so that they reflected the judged magnitudes of his/her experiences. In magnitude estimation, each stimulus was assigned a number that reflected its distance as a proportion of that to the standard. The standard stimulus was set as “100.” If a test stimulus was subjectively twice as far as the standard, a participant was required to assign it a magnitude of “200.” Under the three device conditions, each participant performed five trials for each of the five stimuli.

Analyses: For each ETA device, the product-moment correlation coefficient (r) between the presented distance (i.e., the stimulus actually presented) and the estimated distance was computed. Additionally, the determination coefficient was computed based on linear regression analysis.

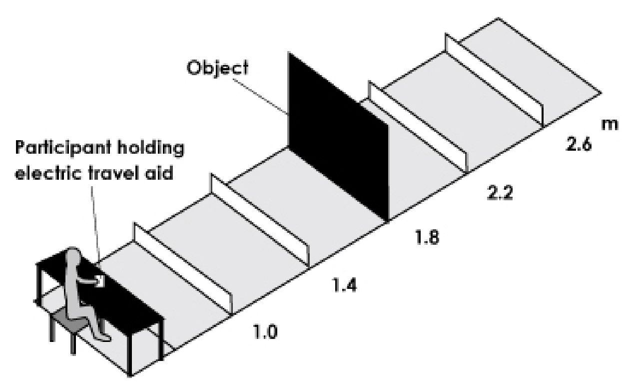



Figure 4. Experimental setup

### C. Result

Figure 5 shows the correlations between the presented and estimated distances measured by each device. As the lines can be seen in Figure 5, the product-moment correlation coefficients of the FB-F, Vibratory device, and Sonar device were 0.917, 0.498, and 0.396, respectively. This indicates that the FB-F provided participants with the most accurate estimation of distance to the presented stimuli.

Figure 6 shows the mean determination coefficients of each group in the three device conditions. A two-way analysis of variance was conducted with group (blind and sighted) and device condition (FB-F, Vibratory device, Sonar device) as the between- and within-subject factors, respectively. The main effect of device condition was significant ($F(2, 28) = 66.68$, $p < 0.01$). Multiple comparison tests between the three device conditions showed significant differences between the devices. By contrast, there was no significant difference between the blind and the sighted groups.

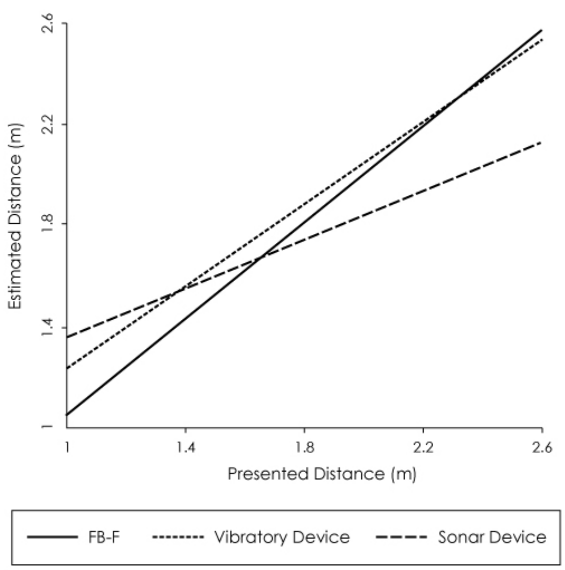



Figure 5. Correlation between estimated and presented distance (n = 400)

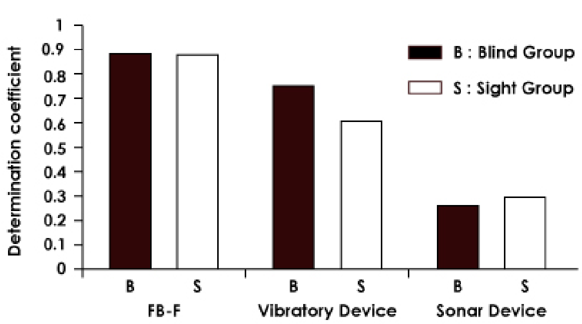



Figure 6. Mean determination coefficients in three devices for the blind and sighted groups

## V. CONCLUSION AND FUTURE WORK

ETAs have three requirements. First, they should be portable and one-handed so that users can carry them easily while locomoting. Second, ETAs should be able to detect the distance and direction to surrounding objects accurately. Such spatial information will allow users to make decisions such as avoiding collisions with obstacles and approaching objects. Third, they should be manageable by anyone regardless of age or cognitive ability so as to provide users with information on the direction of or distance to objects intuitively or directly. In this study, we ascertain that the FB-F satisfies these three requirements from the results of a psychological experiment. To enhance the usability of the FB-F, we will continue experimental studies to evaluate its efficiency in traveling. Specifically, we are conducting a psychological experiment in which the FB-F allows blind persons to reach a destination by using an experimental maze. We will make necessary improvements by evaluating the results collected in experimental settings. We hope that our device, which realizes the idea of a “Future Body,” will allow both visually impaired and healthy persons to extend their capabilities for exploring surroundings.

## ACKNOWLEDGMENT

We wish to extend our special thanks to those who volunteered to participate in our experiment. Our research on the development of the Future Body-Finger was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports and Technology, Japan.

## REFERENCES

[1] A. Clark and D. Chalmers, “The extended mind,” in Analysis, vol. 58, pp. 7-19, 1998.

[2] B. B. Blasch, W. R. Wiener, and R. L. Welsh, Foundations of Orientation and Mobility, 2nd ed. New York: AFB Press, 1997.

[3] C. Carter and K. A. Ferrell, The Implementation of Sonicguide with Visually Impaired Infants and School Children, Sensory Aids Corporation, D. W. Campbell, Ed. Ill: Bensenville, 1980.

[4] L. Kay, “A sonar aid to enhance spatial perception of the blind: engineering design and evaluation,” in Radio and Electronic Engineer, vol. 44, no. 11, pp. 605-627, 1974.

[5] L. Kay, “Acoustic coupling to the ears in binaural sensory aids,” Journal of Visual Impairment & Blindness, pp. 12-16, 1984.

[6] L. A. Johnson and C. M. Higgins, “A navigation aid for the blind using tactile-visual sensory substitution,” in Proc. EMBS ’06, IEEE 2006, vol. 6, pp. 6289-6292.

[7] S. Runeson, “On the possibility of 'smart' perceptual mechanisms,” in Scandinavian Journal of Psychology, vol. 18, pp. 172-179, 1977

# User-Trained, Zero-Configuration, Self-Adaptive Opportunistic Wi-Fi Localization for Room-Level Accuracy

Jef Lathouwers
Maarten Weyn
Charles Vercauteren
*dept. of Applied Engineering: Electronics-ICT*
*Artesis University College*
*Antwerp, Belgium*
*Email: jef.lathouwers@student.artesis.be*
$\{$*maarten.weyn,charles.vercauteren*$\}$*@artesis.be*

***Abstract*—In this paper the possibility of room-level localization through Wi-Fi by using user collaboration and zero-configuration is investigated. User collaboration and zero-configuration means that it avoids the time-consuming training phase known by other systems such as fingerprinting and entering the floor plan. Fingerprints are created as soon as the users start to collaborate by providing their location and corresponding Wi-Fi data. A floor plan is not necessary as fingerprints are simply assigned to rooms without using coordinates. It is called opportunistic localization in the way that it relies on the already available infrastructure, thus no additional hardware needs to be installed. Using these methods, simulations show that a localization success rate of about 90% can be reached and that the system is able to cope with collaboration errors and a changed environment.**

***Keywords*-Positioning Systems; Indoor Wi-Fi Localization; Zero-Configuration; User Collaboration; Adaptive Algorithms.**

## I. Introduction

Due to the popularity of mobile devices with a wireless interface and the increasing presence of wireless access points, indoor localization using Wi-Fi signal strengths is becoming more and more popular. The knowledge of the location of persons and assets is essential for context-aware computing in ambient intelligent environments.

A common technique that utilizes these signal strengths is pattern matching. This approach of matching radio frequency patterns, which are called fingerprints, was initially proposed by Bahl and Padmanabhan [1] in relation to their system RADAR. However, most current systems that apply fingerprinting require a so-called offline phase where the fingerprints are constructed before localization in the online phase can be performed. In this case, a fingerprint consists of the received signal strength of all received access points combined with the current position of the measurement.

Such systems have some implications. First of all, these systems rely on the data gathered in the offline training phase where the properties of the environment can be different than in the online localization phase. A changing environment can change the radio signal strength pattern dramatically because of multipath and shadowing. In addition, Wi-Fi access points can be replaced, removed or added also resulting in a different radio signal strength pattern. Consequently, the fingerprints become inaccurate and recalibration of the fingerprints is required. Secondly, making a radio map of the environment requires the environment to be known, thus a floor plan is needed.

In this paper, the following research question is addressed:

*How is it possible to achieve room-level accurate localization using Wi-Fi through user collaboration and zero-configuration?*

A system is proposed, similar to Redpin [2], which does not require a training phase. In addition, it does not initially need any information about the environment and adapts to changes in the environment. It is called opportunistic localization as it relies on the existing infrastructure. Hence no additional hardware needs to be installed. The key concepts for achieving this are user collaboration and a room-level approach. The system lets the users create and update the locations and fingerprints in order to achieve room-level accuracy. This is accomplished by giving the users the ability to tell the system where they are. Locations are identified by a name and ID and fingerprints are simply assigned to these IDs. No coordinates are used.

However, new problems arise when using user collaboration and zero-configuration. User collaboration comes along with a security issue as users can mislead the system. This can be accidentally or on purpose. As the system completely relies on the input of the users to build fingerprints and aims to be adaptive to the environment, it is hard to distinguish between collaboration errors or a changed environment. Furthermore, because neither a floor plan or coordinates are used, it is hard to link rooms as neighbors. When the system is able to detect motion patterns between measurements, the localization can be improved.

Although the general idea is the same as in Redpin, the

 **ISBN: 978-1-61208-235-6** 

implementation is quite different. Firstly, Redpin focuses on mobile phones and tries to combine Bluetooth with GSM and Wi-Fi signal strengths. In contrast, our system aims to achieve the same results by only using Wi-Fi signals and using a probabilistic approach with a motion and measurement model. Secondly, Redpin relies on the fact that users always collaborate correctly, which will not be the case in real life.

The remainder of this paper is organized as follows. In Section 2 the architecture of the system is presented. Section 3 and 4 go more deeply into the fingerprinting process. Section 5 covers the results of experiments conducted at our Department of Applied Engineering and the conclusions are presented in Section 6.

## II. Architecture

Figure 1 shows the architecture of the system. It consists of two basic components: a client application acting as a sniffer tool that gathers information about all Wi-Fi access points in range, and a server that stores fingerprints in a database and runs a localization algorithm. The system can be divided into three parts that are described below: user collaboration, fingerprinting and localization.

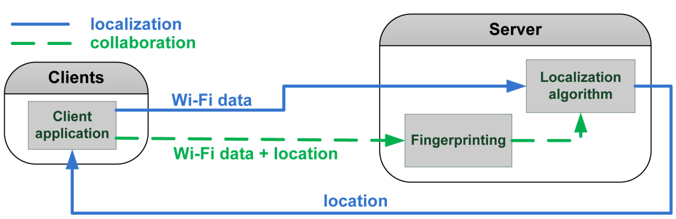



Figure 1. Architecture

### A. User Collaboration

As noted before, fingerprints are constructed based on user collaboration. This consists of two tasks. Firstly, as the system initially does not know anything about any location, rooms need to be added to the system. This can easily be done in the client application by adding the name of the room. Next, fingerprints need to be constructed in order to train the localization algorithm. When connecting to the server a user can tell the system where he or she is by selecting the right location from a list. This is called a collaboration and is represented in Figure 1 by the dotted line. The system starts to perform localization as soon as there is one room added and one collaboration is made. Of course, the system will perform better as more rooms are added and more collaborations are made.

As our system aims to be able to cope with user collaboration errors, each collaboration will be rated and this rating will affect the impact of this collaboration to the fingerprints.

### B. Fingerprinting

Fingerprinting for room-level accuracy needs a special approach. As rooms can be large, measurements from different positions in the same room can be totally different. Averaging all those measurements will lead to an artificial fingerprint. Thus, multiple fingerprints per room are needed. This is accomplished by clustering the different measurements. Data clustering is the process of grouping together similar multi-dimensional data vectors into a number of clusters. The multi-dimensional data in our case is the Wi-Fi measurement that is sent by the user in combination with a location (i.e., a collaboration). Each cluster will eventually lead to one fingerprint.

*1) K-means Clustering:* A widely used clustering method is k-means clustering [3]. Because there is no prior knowledge from training data, it is called unsupervised clustering. It aims to partition $n$ observations into $k$ clusters where each observation belongs to the cluster with the nearest mean. This mean is called the centroid of the cluster and will become a fingerprint in our case. Therefore a measure of distance is needed to compare observations (i.e., Wi-Fi measurements) to each other. A simplified version of a measurement model proposed by Weyn in [4] is used to accomplish this. It is developed to calculate the similarity between a measurement and a fingerprint. Besides the comparison between matching access point measurements, it takes into account, and also makes a difference between, an access point measurement from the fingerprint that is missing in the measurement and an access point measurement that is missing in the fingerprint in comparison to the measurement. This is not relevant when comparing two measurements in between and thus is not used. The algorithm returns a probability $p$, which represents the likelihood of similarity between two observations. The distance between those two observations is defined as $-log(p)$.

*2) Adaptive:* As noted before, fingerprints are constructed from the centroids of the different clusters. Also stated before, each collaboration will have a rating. This rating represents the likelihood that the collaboration is a correct one and affects the impact of that collaboration to the centroid. Therefore a weighted mean is used in the clustering process instead of a normal mean. The weight will not only be the collaboration rating but also includes a time and user rating as in Equation 1:

$$W_c = R_c^y \cdot R_t \cdot R_u \tag{1}$$

$R_c$ and $R_t$ are parameters representing the collaboration and time rating while $R_u$ represents the user rating. $R_c$ is powered to a value $y$, which will determine the influence of the collaboration rating on $W_c$. This is explained in more detail in Section 5.

These three ratings all have a different goal. The collaboration rating will filter collaboration errors, the time

 ISBN: 978-1-61208-235-6 

rating will make the system adaptive to the environment and the user rating will filter collaborations from users that consistently collaborate faulty. These are called bad users.

The collaboration rating is calculated using the standard score of the collaboration. The standard score of a sample is the distance of a sample to the mean of the distribution divided by the standard deviation of the distribution [5]. As a result, it is possible to compare samples from different distributions and thus to rate collaborations belonging to different clusters. The collaboration rating is equal to $2p$ where $p$ is the probability that a random sample has a greater standard score than the one being rated.

The time rating is calculated using Equation 2:

$$R_t = \begin{cases} 1 - \frac{1}{a}x & \forall x \in [1, a] \\ 0 & \forall x \notin [1, a] \end{cases} \tag{2}$$

All collaborations from a room are, starting with the most recent one, ordered in time and given a sequence number $x$. The linear function in Equation 2 will result in a decreasing time rating for older collaborations ending in a time rating of zero for the $a^{\text{th}}$ collaboration. Consequently, value $a$ determines how many collaborations are included in the sample space.

Equation 3 is used to calculate the user rating:

$$R_{u_1} = \frac{w_1 \cdot R_{u_0} + w_2 \cdot \overline{R_c*}}{w_1 + w_2} \tag{3}$$

The new user rating $R_{u_1}$ is the weighted mean of the previous user rating $R_{u_0}$ and the average collaboration rating from all collaborations from that user during the last $a$ collaborations. This is denoted as $\overline{R_c*}$. When a user rating falls below a threshold $t$, all collaborations from that user are expelled from the sample space.

### C. Localization

Localization is done when a user connects to the server and sends the Wi-Fi data gathered by the client application to the server. A Bayesian approach, as discussed in [6], can be used for computing a probability for each room. This is shown in Equation 4.

$$p(l|o) = \frac{p(o|l) \cdot p(l)}{p(o)} = \alpha \cdot p(o|l) \cdot p(l) \tag{4}$$

$p(o)$ is the prior probability of an observation and can be replaced by a normalizing constant $\alpha$ since it is independent of the location. $p(l)$ is the prior probability for a location $l$ and is calculated by the motion model. This is multiplied by $p(o|l)$, which is the posterior probability of an observation given a location. This is determined by the measurement model.

*1) Motion model:* The motion model can be denoted as in Equation 5.

$$p(l_t) = p(l_t|l_{t-1}) \cdot p(l_{t-1}) \tag{5}$$

$p(l_t)$ is the probability for a location at time $t$, which is the probability of that location at time $t-1$ multiplied by the motion probability of the location at time $t$ given the location at time $t-1$.

These motion probabilities need to be learned automatically. This is accomplished by analyzing users their consecutive measurements to detect transitions between locations. Laplacian smoothing, also known as add-one smoothing [7], is used to calculate motion probabilities out of the number of transitions detected. Using this method there will always be a minimum probability for locations that have no detected transitions. For each combination of locations a probability is calculated using Equation 6.

$$p(l_1|l_0) = \frac{T_{l_1} + 1}{T_{l_x} + L} \tag{6}$$

$p(l_1|l_0)$ is the probability that someone will be at location $l_1$ after being at location $l_0$. The numerator consists of two parts. $T_{l_1}$ represents the number of transitions detected between $l_0$ and $l_1$. This is added to 1, which will prevent a probability of zero in case of zero detected transitions. The denominator is the addition of the total number of transitions counted from location $l_0$ to any other location, represented by $T_{l_x}$, and the number of locations $L$ that exists in the database. Using this equation the sum of probabilities for each room equals one.

The algorithm for the transition detection is shown in Algorithm 1. Because the system will not be 100% reliable in its localization, some filtering is used to detect transitions. $l_t$ represents the calculated location of a certain user at time $t$. At this time $t$, a possible transition from $t-2$ to $t-1$ is investigated. Checking on $l_{t-2}$ and $l_t$ prevents a transition detection when the system returned one aberrant location in between other similar locations.

**Algorithm 1** Transition detection

**if** $l_{t-2} = l_{t-1}$ or $(l_{t-3} \neq l_{t-1}$ and $l_{t-2} \neq l_t)$ **then**
  $Transition_{l_{t-2} \to l_{t-1}}$ detected
**else**
  No transition detected
**end if**

*2) Measurement model:* An algorithm proposed by Weyn in [4] is used for computing the probability between a measurement and a fingerprint. The problem of hardware variance between training and localization devices, as discussed by Tsui *et al.* [8], is also tackled by this algorithm.

 ISBN: 978-1-61208-235-6 

## III. Optimal Number of Clusters

Another issue is the detection of the optimal number of clusters per room. Each cluster will lead to one fingerprint by calculating a weighted average. If too less clusters exist for a large room, averaging those Wi-Fi measurements will result in incorrect fingerprints and higher localization error rates.

By running the k-means algorithm multiple times with a different number of clusters, starting with one, the number of clusters will vary per room if this iteration stops when the average distance from all samples to their centroids does not exceed a threshold $i$. As a result, the higher the value for this parameter is chosen, the less clusters and thus fingerprints are created. Using this method, more fingerprints will be formed in larger rooms or in rooms where the radio frequency patterns differ a lot.

A value of 0.75 for $i$ was chosen for the simulations described in Section 5 as this minimized the localization error rate. This value corresponds to an average of 2.2 fingerprints per room. This test is not executed in other environments so no general conclusion for the optimal value of $i$ can be drawn yet.

## IV. A Changing Environment

Samples are defined as outliers when they are beyond the inner fence of the box plot of the sample space as explained in [5]. Because it is avoidable that collaboration errors become a new cluster and thus a new fingerprint, all collaborations that are marked as outliers are expelled from the sample space when trying to find the optimal number of clusters in the way that is explained above. If the number of clusters is found, the k-means algorithm is run one more time, this time including the outliers. Accordingly, the new centroid of the cluster where the outlier is assigned to will move towards the outlier with the result that the distance from the outlier to the centroid will decrease. This will especially happen if more collaborations similar to the outlier will be added. If this happens there is possibly a change in the environment and because the centroid is moved to the outliers, there is a chance that the former outliers will not be classified as an outlier anymore. This is because all collaborations are rated again after each new collaboration. Consequently, they are added to the sample space for finding the optimal number of clusters and possibly a new cluster will be formed.

## V. Results

The 3rd and 4th floor of the Department of Applied Engineering at the Artesis University College of Antwerp were used as a test-bed to analyze the accuracy of the system. A floor plan can be seen in Figure 2. Our system aims to be both adaptive to the environment and being able to filter collaboration errors, but it will not be possible to achieve maximum results on both goals. As a result, two configurations of the system are proposed. These are called the adaptive configuration and the filter configuration. Each of them will perform better on one of the goals and worse on the other. Extremely put, the adaptive configuration will accept all collaborations whether they fit with the existing fingerprints or not, while the filter configuration will be suspicious to aberrant collaborations. As a result, the adaptive configuration will be sensitive for collaboration errors, but also will adapt faster to a changed environment (e.g., a replaced access point). In contrast, the filter configuration ignores collaboration errors but the downside for this configuration is the slower adaption to a new environment. This distinction is made as one can choose what configuration fits best his requirements.

The difference between both configurations is the value for parameter $y$ in Equation 1, which is the power of the collaboration rating $R_c$. This power defines the influence of the collaboration rating on the collaboration weight in the k-means clustering process. The higher the value for $y$ is chosen, the more the collaboration rating will affect the total collaboration weight and thus the impact of that collaboration to the weighted average of a cluster (i.e., fingerprint). A collaboration with a low collaboration rating will thus affect the fingerprint less in the filter configuration than in the adaptive configuration.

Several simulation tests are performed, each testing a different aspect of our system. Because of the difficulty of simulating motions from users, tests A, B, C and D are performed without using the motion model explained in Section 2. Value $y$ in Equation 1 is chosen 0.8 for the filter configuration and 0.25 for the adaptive configuration, a value of 50 is used for $a$ in Equation 2. These $y$-values are chosen this way in order to show the difference between the two configurations clearly. Choosing them more closely to each other will give similar results, but more moderate. The influence of value $a$ is explained in test D.

### A. Localization Rates

Figure 2 shows the localization success rates for each room. Averaging these rates taking into account the surface area of the rooms, an overall success rate of 89.19% is reached. These results are obtained using the adaptive configuration and during optimal circumstances where collaborations were done correctly and performed at different positions in each room.

### B. Random Collaboration Errors

Figure 3 shows the localization success rates in case of random collaboration errors. The two bars at the left represent the localization success rates in perfect conditions where all collaborations were done correctly. Both configurations reach a localization success rate of about 90%.

The localization success rates in case collaboration errors were made are shown more to the right. The filter config-

 ISBN: 978-1-61208-235-6 

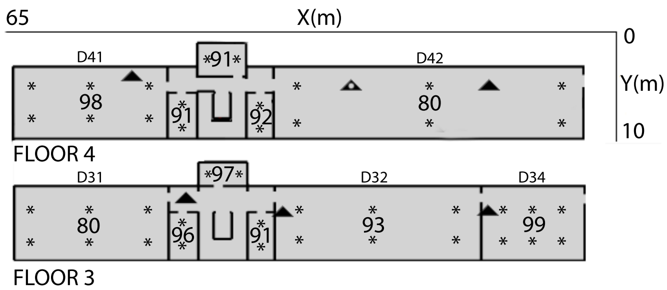



Figure 2. Floor plan Department of Applied Engineering. Access points are represented by a triangle, collaboration points by an asterisk. The numbers in the rooms are the localization success rates.

uration remains more stable than the adaptive configuration when the collaboration error rate increases. This is because the collaboration errors affect the fingerprints more in the adaptive configuration than in the filter configuration. As a result, the fingerprints will move more towards the collaboration errors in the adaptive configuration and the localization success rate will decrease faster in case of collaboration errors. In contrast, the filter configuration is able to filter those collaboration errors better and remains more stable.

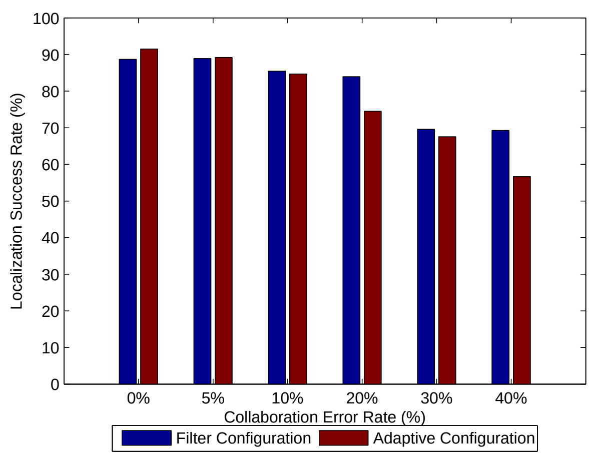



Figure 3. Localization success rates in case of random collaboration errors.

## C. Bad User Collaboration Errors

Figure 4 shows the effect on the localization success rates in case of a bad user. The purpose of the system is to filter his or her collaborations. This is tested by comparing two simulations of the same configuration. Both simulations start with the same fingerprints and during the simulation the same collaboration errors are added. The line with triangles shows the localization success rate if each of those collaboration errors are made by a different user. As more collaboration errors are done, the localization success rate decreases. In contrast, the localization success rate of the line with circles remains stable. The same collaboration errors are now done by only one user (i.e., a bad user). As a result, his or her user rating falls below threshold value $t$, which means all of his or her collaborations are expelled and only the correct localizations from the other users remain in the sample space. Threshold value $t$ is chosen 0.25 during this simulation and for the weighted mean in Equation 3, values 10 and 1 are used for $w_1$ and $w_2$.

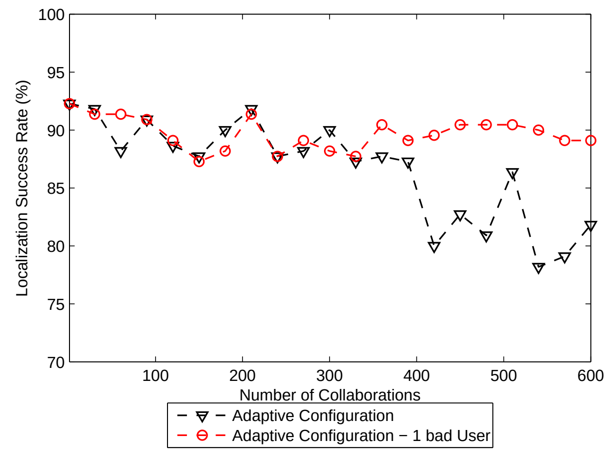



Figure 4. Localization success rates with a collaboration error rate of 20% contributed by 1 user (line with circles) or by multiple users (line with triangles).

## D. Changed Environment

Figure 5 shows the localization success rates in case of a replaced Wi-Fi access point. In our simulation the access point from room D42 that is marked with a white dot in Figure 2 was replaced by another access point after 20 collaborations. The filter configuration will perform worse on this because the collaborations after the change differ from the fingerprints in the database and thus will be filtered. The adaptive configuration reaches the former localization success rate after about 200 new collaborations (i.e., approximately 18 collaborations per room). In contrast, the filter configuration needs about 580 new collaborations to recover. This corresponds to approximately 50 collaborations per room. As a value of 50 for parameter $a$ was used during these simulations, one can conclude that the filter configuration filters all new collaborations until no old collaborations exist in the sample space anymore.

## E. Motion Filtering

The simulation results above were obtained without the use of a motion model. Figure 6 shows the calculated motion probabilities, using the method explained in Section 2, for the rooms of floor 3 after a user walked around the building

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-235-6

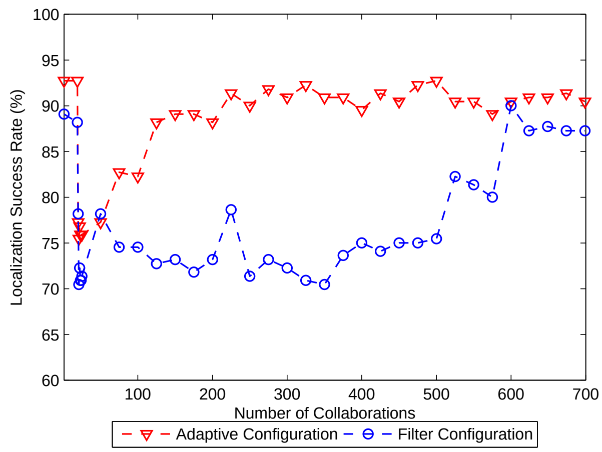



Figure 5. Localization success rates in case of a replaced access point after 20 collaborations.

from Figure 2. The user walked ten times from each room to each of its neighbors, each time staying one minute in each room.

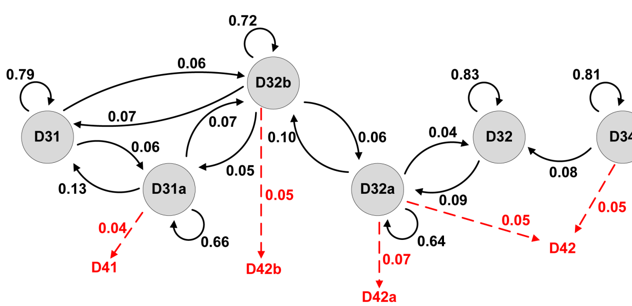
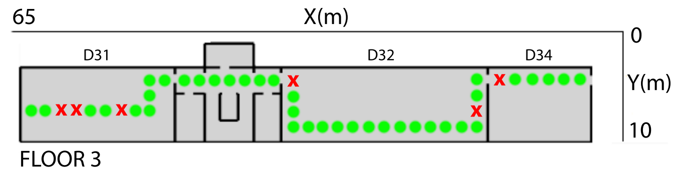



Figure 6. Motion probabilities for the rooms on floor 3. All probabilities smaller than 0.04 are not shown to improve the visibility of the figure.

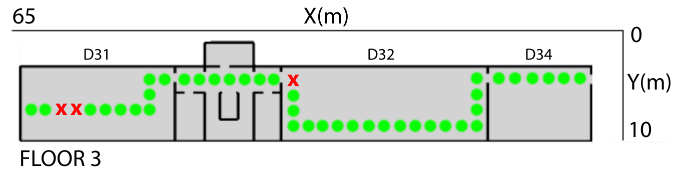
The dotted arrows represent motion probabilities from a room of floor 3 to a room of floor 4, which ideally would not exist if the localization and the transition detection algorithm would be working perfectly.

Using the probabilities from above, a new test was performed simulating a user that walks through the rooms of floor 3 from right to left at a steady pace. The localization results can be seen in Figure 7(a) and Figure 7(b), where a dot and a cross respectively represents a correct and a faulty localization.

Figure 7(a) shows the localization results without the use of the motion model. In contrast, Figure 7(b) shows the localization results with the same measurements, using the motion model with the motion probabilities from Figure 6. As can be seen, three localization errors from the simulation without the motion model are disappeared in the simulation with the use of the motion model. As expected, the motion




(a) Localization test without the motion model.




(b) Localization test using the motion model.

Figure 7. Localization test of a user that walks from right to left. A dot represents a correct localization and a cross represents a localization error.

model acts as a sort of filter on the localization process.

## VI. Conclusion

As a general conclusion, we can state that this research has reached its goals. Using the aforementioned methods, a localization success rate of about 90% is reached on our department in case of 100% correct collaborations. This is comparable to the results of Redpin. Additionally, simulations show the ability of the system to cope with collaboration errors and to be adaptive to a changed environment. Depending on the adaptive nature of the environment and thrust of the users, different parameters to configure or adapt the system are proposed. Lastly, an experiment testing the motion model is presented, which shows the promising ability to filter localization errors.

The proposed systems enables context-aware applications without the need for an offline fingerprinting phase to calibrate the system. It uses the input of the users to create a trustworthy system, which is created during the use of the system itself.

## References

[1] P. Bahl and V. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, 2000.

[2] P. Bolliger, "Redpin – adaptive, zero-configuration indoor localization through user collaboration," in *Proceedings of the First ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environment Computing and Communication Systems*, San Francisco, USA, Sep. 2008.

[3] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.

[4] M. Weyn, "Opportunistic seamless localization," Ph.D. dissertation, University of Antwerp, Mar. 2011.

 ISBN: 978-1-61208-235-6 

[5] N. e-Handbook of Statistical Methods. (2012) http://www.itl.nist.gov/div898/handbook/ [May. 21, 2012].

[6] Widyawan, M. Klepal, and D. Pesch, “A Bayesian approach for RF-based indoor localisation,” in *IEEE ISWCS*. Trondheim, Norway, October 2007.

[7] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2009, p. 863.

[8] A. Tsui, Y. Chuang, and H. Chu, “Unsupervised learning for solving RSS hardware variance problem in WiFi localization,” *Mobile Networks and Applications*, vol. 14, pp. 677–691, 2009.

# Requirements of Task Modeling in Ambient Intelligent Environment

Asma Gharsellaoui,Yacine Bellik
AMI Group, LIMSI-CNRS, Paris-South University
Bât 508, Plateau du Moulon, B.P, 133,
91403 Orsay Cedex, France
asma.gharsellaoui@limsi.fr
yacine.bellik@limsi.fr

Christophe Jacquet
Department of Computer Science,
SUPELEC Systems Sciences,
3 rue Joliot-Curie
91192 Gif-sur-Yvette Cedex, France
Christophe.Jacquet@supelec.fr

***Abstract*— User tasks modeling has been the focus of many research works either for designing user interfaces suitable to given use cases or for identifying the current task of the user in order to assist him/her. The problem of user tasks modeling has been widely addressed within the context of GUIs (Graphical User Interfaces). In this paper, we present an overview of several existing task models and we discuss the specificities of user tasks modeling in the context of Ambient Intelligent Environments, which are mainly place-related tasks, device-related tasks and the possible errors occurring while performing a task. Then, w e discuss the limitations of the classical user tasks models and depict the requirements of a task model specific to an interaction with an ambient environment.**

***Keywords-User Task modeling; Ambient intelligent environments; User-centered systems.***

## I. INTRODUCTION

Ambient Intelligent (AmI) environments strive to support everyday life of the user through their embedded intelligent objects [1]. The goal of such environments is to satisfy user needs based on his/her preferences. Users are the center of such systems and are becoming more and more accustomed to technology and require more complex interaction possibilities. A user interacts with the system through tasks which are actions the user is going to perform in order to achieve a certain goal. Research in the field of assistance and supervision of tasks in ambient environments are still at an early stage.

In this context, our work focuses on helping users to perform tasks properly through an adapted modeling approach. The purpose of user task modeling is to paint a predictive picture of user's activity. This picture must be detailed enough to create usable systems. Section II details the basic concepts related to task modeling. Section III contains a discussion of existing task models and their limitations. Section IV presents the specificities of ambient tasks and is followed by a discussion on how existing models can or cannot deal with these specificities through a scenario example (Section V). Our main contribution lies in Sections VI and VII that, respectively, introduce a new taxonomy for tasks categorization in ambient environments, and identify requirements for task models adapted to ambient applications. Section VIII gives conclusions and directions for future work.

## II. BASIC CONCEPTS

### A. Task

According to the Shorter Oxford Dictionary, a task is "any piece of work that has to be done". Tasks can be seen as actions that have to be performed to reach a goal in a specific application domain. They can be either logical such as retrieving information about a contact or physical such as dialing a phone number [2]. These actions are intended to be performed by the user, but they do not necessarily correspond to the user behavior while manipulating the application.

### B. Goal

A goal is an intention either to modify the state of an artifact or to maintain it [3]. For example, accessing a contact database to look up a phone number of a colleague is a goal which does not require the modification of the state of the application, whereas accessing a contact database to add a new contact requires the modification of the state of the application.

Tasks and goals are closely connected. Each task can be associated with one goal that is the goal achieved by performing the task while one goal can be achieved by performing one or multiple tasks. In some cases it is possible to choose among different tasks to achieve a certain goal [2].

### C. Activity

A user activity is a set of actions that the user is really going to perform when interacting with the system to achieve a goal. The activity corresponds to the user behavior when interacting with the system depending on the context of use. It is unpredictable and obtained by observing the behavior of the user.

### D. Task Model

When designing an interactive system, several phases are conducted. First of all the task analysis phase [4] takes place to identify the fundamental elements of a job and to examine required knowledge and skills [5]. When this phase is finished, relevant tasks are identified and the task modeling phase starts. The purpose of task modeling is to build a model which precisely describes the relationships among the various tasks already identified. A task model describes the predictive activities to be performed in order to reach user goals.

The basic principle of task modeling is to decompose each task until reaching basic tasks which cannot be further decomposed. In some cases basic tasks require one single physical action to be performed. The level of decomposition in task modeling depends on its purpose. Task models should be rich in information and flexible so as to capture all the main activities that should be performed to reach the desired goals and the different ways to accomplish them [6]. We notice that there is a difference between task modeling and process modeling. A process modeling is a way of describing actions that must be done by the system and so implemented. But task models describe what could be done but not necessarily what will be done.

Task models can be useful for two main classes of people [2]:

- Designers: task models provide a structured approach to evaluate an existing system or as a starting point to design new applications from scratch before implementing them.
- End users: the developer takes into account the user's activities since the physical actions supported by the user interface can be mapped onto logical actions and the representations provided could support the possible tasks.

## III. Existing Task Models and Their Limitations

A number of approaches to task modeling have been developed. In the following paragraphs, we present a short overview of the main ones.

HTA: Its main idea is to represent complex tasks through a hierarchical representation (tree-like structure) [7]. This process starts with the identification of goals [8] knowing that each goal has a status (i.e., latent or active) and conditions to be satisfied. Then the decomposition is repeated recursively until obtaining observable subtasks allocated either to the user or to the user interface [9]. Using this formalism we don't see clearly if we model a system action or a user one. This model doesn't offer a lot of operators. For example there is no specific symbol to model parallel actions; one can only choose to use non-ordered sequences.

GTA: It describes the tasks by focusing on four central concepts: agents, roles, work [7] and situation [10]. Agents often represent people, either individuals or in groups; in some cases there can be non-human actors, or systems that comprise collaboration between human agents and machines. GTA doesn't enable one to see if the task is allocated to the system, the user or both (case of an interactive task) since there is no single representation for each type. It enables one to express only few operators in a specific way, for example hierarchical structure is represented from left to right.

CTT: The Concur Task Tree is a concurrent notation having a structured representation (a tree-like form). There is a distinction between four groups of tasks with various representations: User tasks, Application tasks, Interaction tasks and Abstract tasks (tasks that need to be refined).It provides a large number of temporal relationships between the tasks [11]. It is one of the richest task models, but it is a static model (once established it doesn't change, see Section V). In addition, it does not focus on the physical space where the task will occur or on the devices since it concerns a single system.

UAN: It is a textual notation where the interface is represented as a quasi-hierarchical structure of asynchronous tasks, the sequencing within each task being independent of the others. Each basic task is associated with one table. These tables have three columns indicating the user actions, the system feedback and the internal state modifications using some specific symbols [12]. A rich set of operators to describe temporal relationships among tasks is available. But this method leads to a large specification. And it is not possible to model actions out of the interaction with the interface.

TKS: This method is based on a Task Knowledge Structure, which is a conceptual representation of the knowledge a person has stored in her memory about a particular task [6]. It has a tree like form. TKS focuses on: Roles, Goal structure and Taxonomic structure. Using this formalism it is not possible to model various relations between tasks as there is no special notation for repetitive, optional, parallel actions. Another limitation of this method is that is not possible to model system actions.

DIANE+: DIANE+ formalism [13] is a tree-like model using three types of actions: interactive, automatic and manual [6]. Diane+ can represent all the constraints of the above specifications such as ordered sequences, unordered sequences, loops, required choices, free choices, parallelism, default operations, and so on [7]. It doesn't include any possibility to put additional information about the task concerning, for instance, location, required devices, etc.

TOOD: It consists of an object-oriented method for modeling tasks in the domain of processes control and complex interactive systems [7]. It is based on a hierarchical decomposition into tasks modeled using Petri Nets. This method consists of four steps: hierarchical decomposition of tasks, identification of descriptor objects and world objects, definition of elementary and control tasks and integration of concurrency [14]. This description covers task hierarchy and temporal ordering. Sometimes we have a succession of user actions and we can't model it using this formalism which obliges each action to be between two system transitions.

## IV. Specifities of Tasks in Ambient Intelligent Environments

### A. Ambient Intelligence

The term Ambient Intelligence (AmI) was coined by Philips in 1998, then used by the European Commission in 2001 and has since been adopted worldwide [16]. At the beginning, AmI was a vision of the future in which environments support the people inhabiting them. It has been introduced [15] as an intelligent, embedded, digital environment that is sensitive and responsive to the presence of people.

AmI has a wide range of application since applications in various domains were realized. According to [16] the main application areas are: AmI at home for domestic care of the elderly [17] or assisted living; AmI at Shops, shopping, recommender systems and business and also many applications in Museums, tourism, groups, and institutions.

Unlike a workstation whose actions have an impact only on itself and its immediate environment, an ambient system can incorporate features that enable it to act more broadly on the physical environment (actuators) or to acquire information from its environment (sensors). In addition to the sensors and actuators, the system also incorporates devices enabling the interaction with the user [18]. Nowadays, sensors available in ambient systems may provide clues about the behavior of people (position, orientation, displacement, etc.). This kind of information could then be exploited to ensure the proper conduct of tasks performed by the user and to offer assistance when needed.

### *B. Characteristics of tasks in AmI Environment*

**Device-dependent tasks**

Ambient systems include many small devices in the surrounding space of the user. Fulfillment of a task may require resources from the environment in which it will take place. That is why we talk about device-related tasks.

If a task requires a specific device, we must find this equipment. In an AmI environment, the devices could be added or removed at each moment or could be used by another person. When the user needs a device, the system must invoke this particular device or an equivalent one. This requires devices' categorization according to the services it could offer to the user. For example, the screen laptop could be used as a projection screen for viewing a film if the TV is out of order or being used by another person. So for each device we will have a list of possible related tasks.

**Place-dependent tasks**

Some tasks may need some resources which may be not mobile but located somewhere in the physical space. The execution of the related tasks will be possible only in a certain range of the physical space. That's called place-related tasks. For example: Kitchen-related tasks are the tasks possible in the kitchen such as: make tea, wash dishes, and prepare dinner...

Tasks in AmI environments also share some characteristics with classical tasks, in particular:

**Temporal relationships**

Two (or more) tasks in an ambient environment could be associated through kinds of relationships. Parallel tasks should be executed at the same time. Sequential tasks need to be executed in order: the beginning of a task corresponds to the end of the previous one. In some cases we could have the choice to get into a state by executing a specific task or another that is equivalent. In other cases we may need to model a second task that suspends the execution of a first task and that will be executed again after the second task finishes.

**Error-prone behavior**

Since we model user actions in interaction with various intelligent embedded systems, it is possible that errors occur during the process (system errors or user manipulation errors). These errors must not be neglected during the modeling phase and alternative solutions should be provided.

## V. DISCUSSION OF EXISTING TASK MODELS THROUGH AN EXAMPLE

A large number of task models have been developed, especially in the context of GUIs. In this section some of these models are discussed according to the characteristics of tasks in AmI environments through an example scenario.

### *A. Ambient scenario 1 "Renewing passport"*

Elie was browsing traveling web pages. The system understands that she plans to go on a trip and reminds her that her passport is no longer valid. The system asks her if she wants help to renew her passport. Once she validates this option, the system starts giving her instructions.

The system connects to the "city hall" website (of the town where she lives) to get the different documents that should be provided when renewing a passport.

The system instructs her that she needs a certificate of birth. She asks for help since she forgets its place. The system mentions that it is in the shelf where she puts her papers. Elie finds it and presents it to the system scanner. The system scans the found certificate and notices that it is expired (it was provided since more than three months) and a new one must be extracted from the "town hall" where she was born. For this purpose, the system suggests Elie to send them an email. An email is automatically written including some information about her (first name, last name, date and place of birth …) asking for a new certificate of birth. Elie validates the email which is sent to the "town hall".

Now, the system indicates that she needs a new normalized photo. Elie goes in front of the camera. First she is in front of a white wall; the photo web service indicates that the photo is not valid since the background must be clear but not white. The second trial is in front of a black wall and the photo system rejected this picture too. The third one is accepted by the system and then printed. Elie takes the pictures and puts them on a tray where she collects all the needed items.

Then the system instructs her to buy a tax stamp from the web site of the general treasury. The system opens the website of the general treasury and Elie chooses the appropriate stamp indicated on the web site of the "city hall" then proceeds to the payment. She types her credit card number and prints the stamp received by email. She adds it to the tray containing all the other papers. She is not sure that she bought the correct stamp so she scans its barcode using a tablet computer and the system tells her that it is indeed the asked one.

At this stage, she needs to find her old passport; she asked the system to help her to find it because she forgot where she put it. The system instructs her to look in the nightstand. She finds it, brings it and adds it to the tray.

Now, that she has collected all the papers she must wait for the certificate of birth to have the file closed and to go to the "city hall" to deposit the renewing papers. However she decides to already get an appointment to deposit the documents.

 ISBN: 978-1-61208-235-6 

The system dials the number of the "city hall" in order to make an appointment. When Elie talks on the phone, her calendar is displayed starting from next week (as the certificate takes usually one week to be delivered) in order to see her free times to insure her availability for the chosen date.

### *B. Task models representation*

In Section 3 we have talked about seven task models. We have tried to model the proposed scenario using some of them to highlight their limits.

Figures 1 and 2 show the modeling of this scenario using the HTA task model. As it could be seen from the model, we couldn't model system tasks so we feel that a big part of the model is missing (TKS model exhibits the same limitation). For example, we don't see that the system proposes help. At first we see directly the user validating help request. We could not model the parallel actions "talking on the phone" and "seeing the calendar". We could not model the possible errors while taking photos directly but we represent them as a repetitive task.

Figure 3 shows a part from the renewing of the passport scenario model using the GTA task model. We can see that we could model system actions as well as user actions but we could not differentiate between the two types since we use the same representation. At the starting of each subtasks we put the time relationship between them and we have various options for that (for example we have "seq" for sequential tasks, "and" for parallel executed tasks, etc.).

Using UAN, we cannot model actions out of the interaction with the interface so we couldn't model the scenario since it includes many actions of an interaction with an ambient system not necessarily a GUI. We could just model for example the purchase of the stamp as it is realized on line.

Figure 4 shows a part of the scenario modeled using the CTT task model. We can see that this model enables to model user tasks, system tasks and interaction tasks with different representations; but, we need to differentiate between different sub-systems since the user is interacting with different sub-systems in order to achieve his/her goal. We also need sometimes to put constraints related to places for instance in our scenario the user must sit in front of the camera to be able to take the picture. This model is rich with its different time operators but once established it doesn't evolve according to the context (when modeling ambient tasks, the task model must evolve in real time to ensure that it takes into account the modification in the environment). There is no special notation for errors occurring.

### *C. Discussion*

Only few of the existing task models include a clear differentiation in the representation of the tasks between system tasks, user and interaction ones. For instance from the ones cited in the third section only CTT and DIANE+ explicitly introduce several categories, whereas the others represent all task types on the same way. None of the referred task models has a special notation for device-or-place-related tasks. This limitation is due to the fact that these task models deal only with tasks invoking only one system, so there was no need to model several systems or place-related systems. Some of the listed task models (for instance: GTA, CTT UAN...) have a large choice of time operators offering possibilities to express various time relationships between tasks. It's also important to notice that not all the task models have a graphical presentation for example UAN doesn't have a graphical representation but just a textual description of the different tasks.

Table I summarizes the main characteristics of existing task models. The column "Types" indicates if the model includes any differentiation between task categories. "Devices" and "Places" show if there is any possibility to express device-related and place-related tasks. "Time Operators" refers to the presence of time operators to express temporal relationships between tasks. The final column shows if there is any graphical representation of the model.

TABLE I. COMPARISON OF EXISTING TASK MODELS CHARACTERISTICS

| | **Types** | **Devices** | **Places** | **Time Operators** | **Graphical Representation** |
|---|---|---|---|---|---|
| HTA | | | | x | X |
| GTA | | | | | X |
| CTT | x | | | x | X |
| UAN | | | | x | |
| TKS | | | | x | X |
| DIANE+ | x | | | x | X |
| TOOD | | | | | X |

## VI. PROPOSAL OF A TAXONOMY OF TASKS IN AMBIENT ENVIRONMENT

Tasks in AmI environments have some specific characteristics; let us try to depict them based on a second scenario.

### *A. Scenario 2*

Elie goes to work at 8am, so she leaves home; the heater is turned off since nobody is at home. She comes home after 6pm so she planned a TV program recording at 5pm for her favorite film. The system starts also another recording of a cooking TV program since it knows that she likes that kind of program. When she leaves the office, the system puts the heater on, 30 minutes before she arrives (after taking into account the distance between the house and the office and depending on the traffic) as the house takes about 30 minutes at least for warming. When she comes back home, she calls her friend John and at the same time turns the TV on and asks for the second channel. Since the TV is in front of the window, the system decides to close up the window stores because the sun shines outside, disturbing watching TV.

 ISBN: 978-1-61208-235-6 

TABLE II. TASK CATEGORIES THROUGH EXAMPLES

| | automatic | implicit interactive | explicit interactive | User |
|---|---|---|---|---|
| Programmed | Recording TV programs programmed by the user | Put heater on when the user comes back home | Record TV program immediately after the user asks for that. | The user asks the system to remind him/her or them to do a certain task at a specific time. |
| Learned | Regular TV programs recording learned from user habits | Put heater on when the user comes back home because the user asked for that many times so the system learned this. | Modification of the intended/planned action after interactions with the user. The system learns how to respond to a new order. | The system notices that each time the user puts the conditioner on, he closes the windows. This task must be added to the task model to enable reminding the user to do it. |
| Deduced | Recording a given TV program since it is similar to what the user is used to watch. | Put on the heater when the user leaves the office (deduced in fonction of the traffic and the distance between home and the office). | When the user turns on the TV, the system closes the windows as it disturbs TV viewing. | The system has meta-rules saying that if the windows are opened the conditioner must be closed. If the user opens the windows and could not turn off the conditioner as the remote controller is disabled,the system asks the user to put off the conditioner manually. |

### B. *New tasks categorization*

Starting from the Scenario 2 and trying to imagine other possible situations, we propose a new categorization of tasks in ambient environments. Tasks are categorized with respect to 2 axes:

- **The way the task is launched**: Tasks can be automatic, explicit interactive, implicit interactive, or done by the user.

  1/*Automatic tasks:* the system does the task by itself without user intervention at this moment.

  2/*Interactive explicit:* the user action or situation is understood by the system as an entry condition or order to a certain task.

  3/*Interactive implicit:* the user directly instructs the system to do a certain task.

  4/*User:* the task is done by the user.

  For categories 2 and 3, there is an interaction between the system and the user which leads to the realization of the task.

- **The way the task is acquired**: Tasks can be programmed, learned or deduced, expressing the way the system gets knowledge of the task.

  1/*Programmed:* the user or the designer instructs the system to do a certain task.

  2/*Learned:* the repetitive invocations done by the user could be learned by the system and then anticipated and realized without user invocation. Here the system learns user habits or repetitive tasks and starts doing them without being ordered explicitly to do that.

  3/*Deduced:* The intelligence of the system includes some rules that allow some tasks to be initiated by the system to satisfy potential needs of the user, based on his/her profile.

We propose to categorize a task through the intersection of these two dimensions, depending on the way a certain task is launched and acquired. Table 2 explains each type through an example from the Scenario 2 or a variation of it.

## VII. REQUIREMENT OF TASK MODELING IN AMBIENT INTELLIGENT ENVIRONMENT

From the previous scenarios, we can extract the following requirements for an ambient task model.

**Place-related and device-related tasks**

Concerning the constraints of place-related and device-related tasks, the task model should include the possibility of labeling tasks with spatial constraints and devices involved to achieve the task. However it is impractical to refer to a specific device in an AmI environment because in AmI environments devices and services may appear and disappear dynamically. So instead we should refer to a family or type of device.

The usage of an ontology for device classification is necessary. One solution is to deal with services which could be realized by using different facilities from the surrounding environment and so the devices will be part of a category offering certain services. Once a task needing a given missing device starts, the system may propose to realize it with one of the available equivalent ones. For example if the user wants to call someone and at this moment the phone's battery is discharged, the system could opt for an equivalent solution which is to call on Skype using the user laptop, So these two solutions could be considered as equivalent since they enable the same goal in two definitely different ways to be achieved. Such devices classification was realized in previous research works; we can cite for example DogOnt, an ontology for ambient intelligent environments [19].

**Granularity or task decomposition**

The granularity of task decomposition must be chosen. Since the basic principle of task models is to decompose a complex task into smaller subtasks until reaching a basic task which could not be further decomposed, the level of decomposition at which we stop decomposing a task in an ambient environment must be determined. We propose to stop the decomposition at the level where services are invoked.

**User profile**
Ambient intelligent environments may be proactive systems. They evolve according to the user profile in order to anticipate his/her expectations. So the model must integrate a part in which user habits and preferences are continuously stored.

**Dynamic model**
The proposed task model should be an evolving model in order to respond to the continuous changes of the context and to enable the representation of the learned and inferred tasks. In contrast to static classical task models, a task model adapted to AmI environments is dynamically changing with the new learned knowledge. It evolves once the system learns more about the user profile or new tasks. This task model should be able to model the reactive part of the tasks as well as the proactive one.

**Error-prone behavior**
Errors may occur during the interaction with the system so the model should have a special notation for either system or user errors and should be able to propose another alternative to overcome errors, for example by a system action or by proposing another alternative to the user. In ambient environments, error detection can go further since we have sensors enabling to detect any equipment's failure to overcome.

**Task categorization**
An ambient task model should refer to the new proposed task categorization by offering a means to label tasks according to their categories. The usage of this taxonomy will be helpful in many cases. For example when we have concurrent tasks to be performed using the same device we wish to prioritize them. Namely deduced tasks would have the least priority since they are only based on the system deduction whereas programmed ones should have greater priority.

Some of the existing task models offer facilities that should be offered by an ambient task model as well. For instance a model having a graphical structure is easy to use by both developers and end users because it highlights the main information and keeps out of sight complex information. A hierarchical structure could also be helpful to enable a clear identification of task steps and to reduce the complexity of the tasks that should be performed in order to reach a certain goal.

The model should also include a rich set of temporal operators to express any possible temporal relation between tasks.

## *D.* Conclusion And Future Work

Task modeling is an interdisciplinary research area which requires knowledge in computer and cognitive sciences and Human Computer Interaction (HCI). In cognitive sciences there is a big focus on how to characterize and identify relevant tasks. In computer science the focus is much more on finding notations suitable to represent tasks and their relationships more precisely. The last area involved is HCI since we try to study and to model a human in interaction with a system.

In this paper, we have made an overview of the existing task models and we have studied their limitations when used to model ambient scenarios.

Ambient tasks share some characteristics with classical tasks but they also have proper specificities. We have detailed the characteristics of tasks in AmI environment. Since the tasks may invoke specific devices in specific places, this has led us to specify device-related and place related-tasks.

We have proposed a new taxonomy of tasks adapted to ambient environments, based on two axes: the way tasks are launched and the way they are acquired. We have given precise examples for each possible category of this taxonomy.

From the previously detailed characteristics, we have deduced a set of requirements for a task model suited to tasks in an ambient environment. In summary such a model must:

- be an evolving model that takes into account changes that may occur in the surrounding environment,
- have services as the lowest level of task decomposition,
- give special labels to each task to specify if the realization needs special devices or must take place in a special area,
- propose a type for each category of tasks (from the proposed taxonomy),
- explicitly represent possible errors and solutions,
- offer a large set of temporal operators.

At this stage relevant characteristics of tasks in ambient environments are highlighted, new task categorization was defined and the requirements of a task model adapted to AmI environment are clearly identified. As a future work, we plan to develop this novel ambient task model that will enable an intelligent assistance to users in their daily activities. This task model will be exploited in real time in order to compare the tasks actually performed with the intended ones and to infer the next actions of the user and adapt system responses. We also plan to have real scale experimentations in an equipped smart room at our Lab (LIMSI-CNRS) based on different scenarios.

## References

[1] G. Riva, M. Lunghi, F. Vatalaro, and F. Davide, "Presence 2010: The Emergence of Ambient Intelligence", in: G. Riva, W.A IJsselsteijn (Eds.), Being There: Concepts, effects and measurement of user presence in synthetic environments, IOS Press, Amsterdam, 2003, pp. 60-81.

[2] F. Paterno, "Task models in interactive software systems", Handbook of Software Engineering and Knowledge Engineering, Vol 1, 2002, Publisher: World Scientific, pp. 1-19, ISBN: 981024973X, DOI: 10.1002/9780470757079.ch4.

[3] E. Schulungbaum, "Model-based user interface software tools current state of declarative models", in the Graphics, Visualization and Usability Center' technical report in Georgia institute of technology GIT-GVU-96-30 November 1996.

[4] J. Annett and K.D. Duncan , "Task analysis and training design", Occupational Psychology, 41, 1967, pp.211–227.

[5] T.C. Ormerod, A. Shepherd, "Using Task Analysis for Information Requirements Specification: The Sub-Goal Template (SGT) Method", The Handbook of Task Analysis for Human-Computer Interaction. Mahwah, NJ.: Lawrence Erlbaum Associates, 2004, pp. 347-365.

[6] W3C group,"Task Meta Models", February 2010, Online: http://www.w3.org/2005/Incubator/model-based-ui/wiki/Task_Meta_Models.

[7] J. Guerrero-Garca, J. Vanderdonckt, and J.M. Gonzalez-Calleros. Towards a multi-user interaction meta-model, Working paper 08/28, 2008.

[8] D. Diaper and N. A.Stanton, “The handbook of task analysis for human computer interaction”,2004 pp.67–116.

[9] S.Mills. “Contextualising design: Aspects of using usability context analysis and hierarchical task analysis for software design”, Journal of Behaviour & Information Technology archive, Vol 26, Issue 6, November 2007, pp 499-506.

[10] Gerrit C. van der Veer , Bert F. Lenting , Bas A.J. Bergevoet. “GTA: Groupware task analysis - modeling complexity”, Acta Psychologica, 1996, vol 91, pp.297-322.

[11] P. Rigole, T.Clerckx, Y. Berbers, and K. Coninx. “Task driven automated component deployment for ambient intelligence environment”, Pervasive and Mobile Computing, vol 3, Issue 3,2007, pp.276-299, ISSN: 15741192, DOI: 10.1016/j.pmcj.2007.01.001.

[12] J. Coutaz, F. Paterno, G. Faconti, and L. Nigay. Comparison of approaches for specifying multimodal interactive systems,1994.

[13] J-C. Tarby and M-F. Barthet, “The Diane+ method”. In J. Vanderdonckt, (Ed.), Computer-aided design of user interfaces,1996, pp. 95–120. Namur, Belgium: Presses Universitaires de Namur.

[14] F. Moussa, M. Riahi, C. Kolski, and M. Moalla. Interpreted petri nets used for human-machine dialogue specification,1st ed., vol. 9. IOS Press: Integrated Computer-Aided Engineering, 2002, pp.87–98.

[15] A. Gaggioli, “Optimal experience in ambient intelligence”,in Ambient Intelligence, IOS Press, 2005, http://www.ambientintelligence.org.

[16] F. Sadri. “Ambient intelligence: A survey“. ACM: computer survey, Vol.43, Issue.4, Ocober 2011, pp.36–66, DOI=10.1145/1978802.1978815,http://doi.acm.org/10.1145/1978802.1978815.

[17] F.Sadri, “Ambient Intelligence for Care of the Elderly in Their Homes”. In Proceedings of the 2nd Workshop on Artificial Techniques for Ambient Intelligence (AITAmI), 2007,pp 62-67.

[18] G. González, , A. Cecilio, B. Lopez, , and de la Rosa “Smart user models for ambient recommender systems“. In Ambient Intelligence and (Everyday) Life., University of Basque Country, San Sebastian, Spain, pp. 113-122. , 2005.

[19] D.Banino and F.CornoDogont-Ontology, “Modeling for Inteligent Domotic Environments” ,” Springer-Verlag Berlin Heidelberg 2008, pp. 790-803.

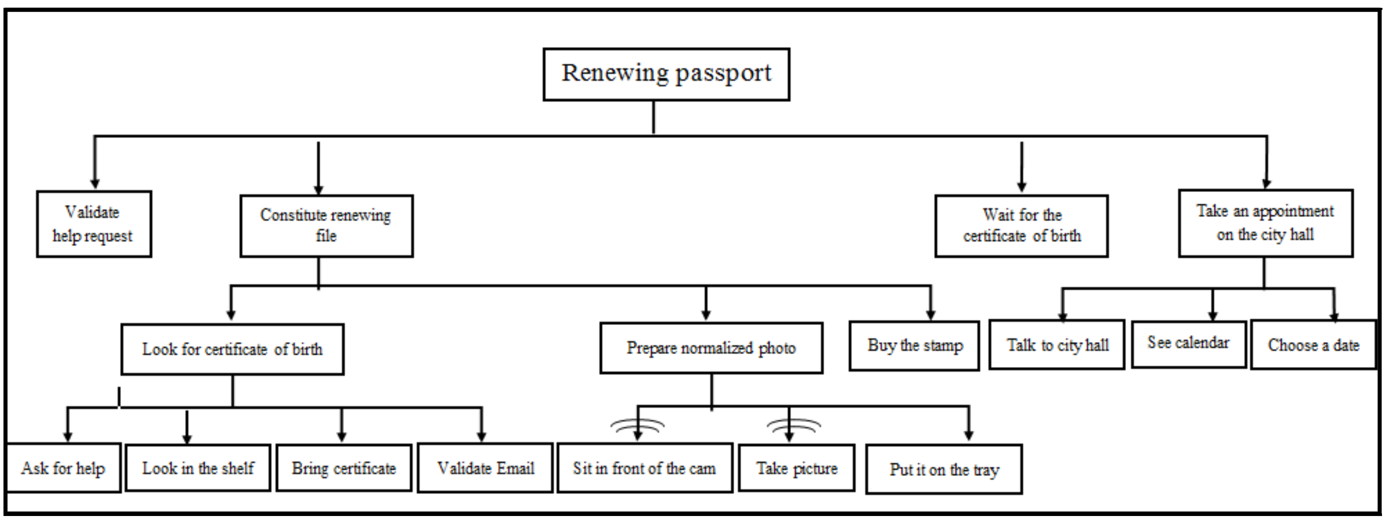




Figure 1. Renewing passport task represented with the HTA task model (part1).

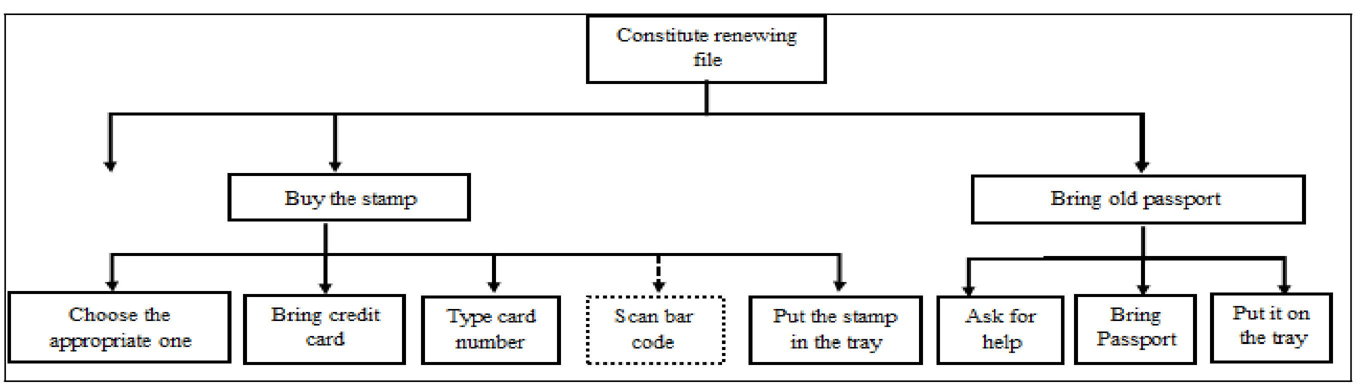



Figure 2. Renewing passport task represented with the HTA task model (part2).

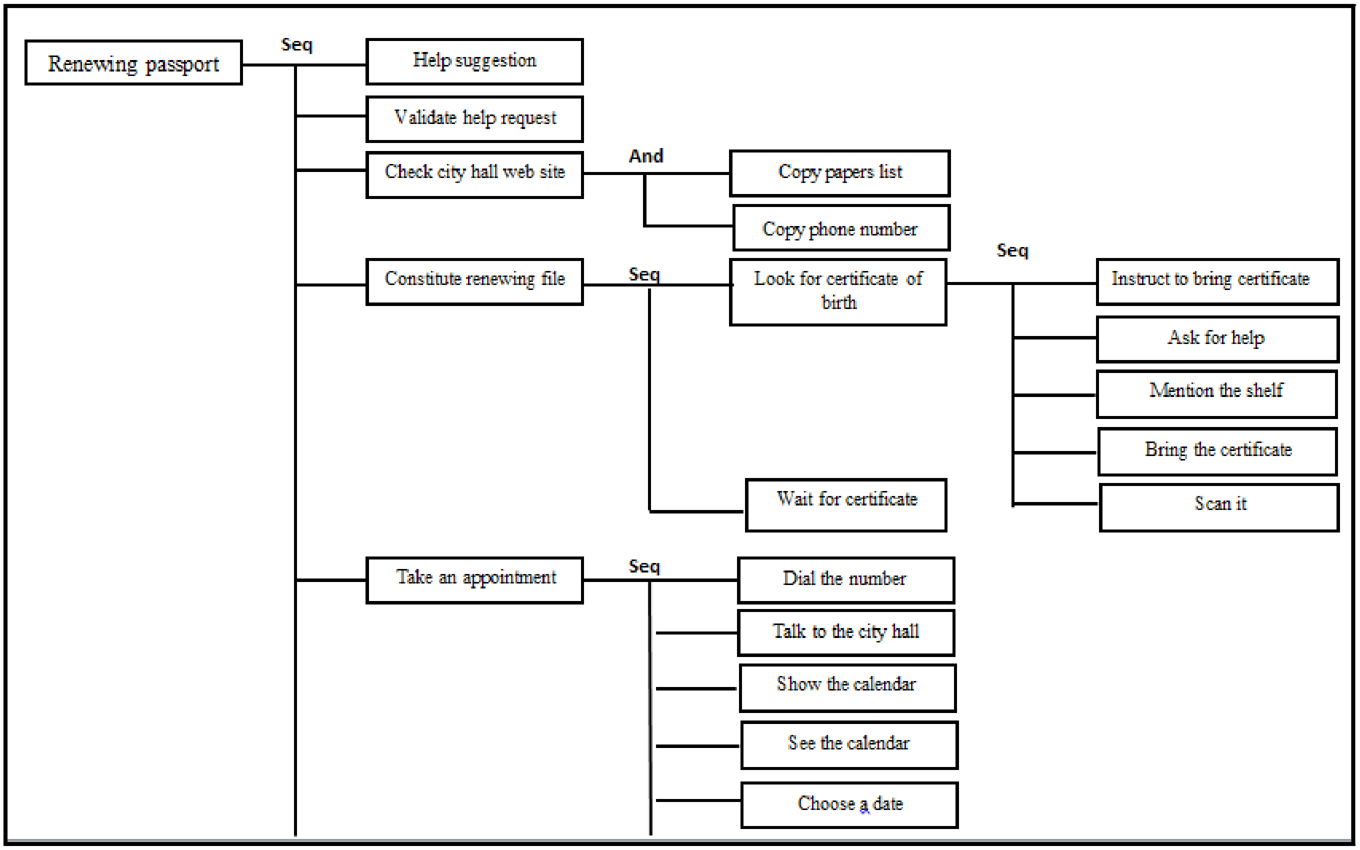



Figure 3. Part from renewing passport task represented with the GTA task model.

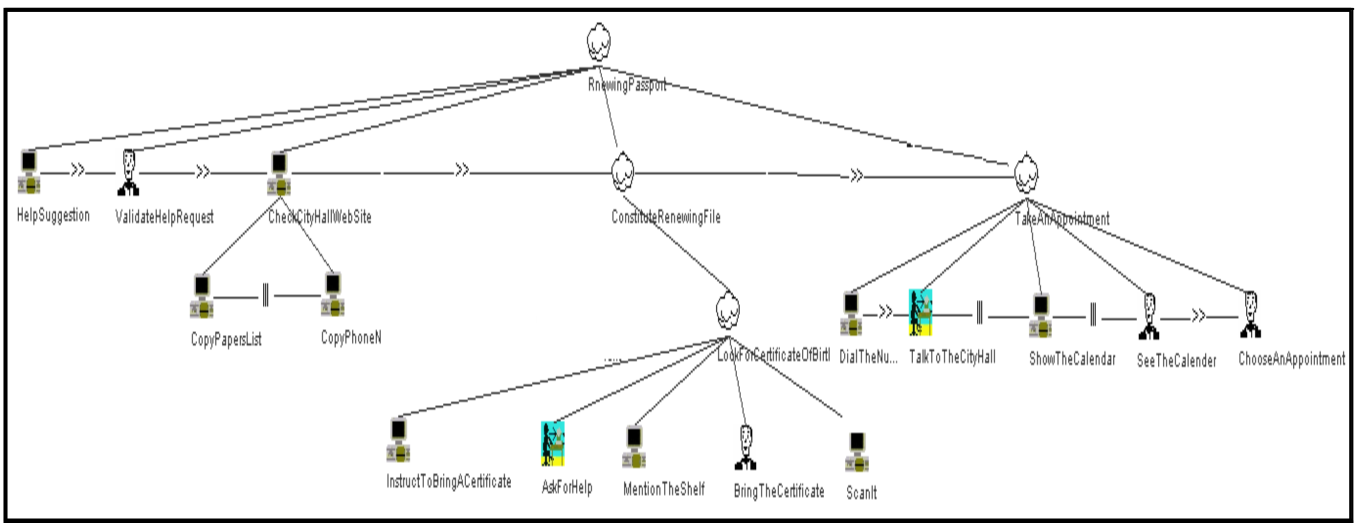



Figure 4. Part from renewing passport task represented with the CTT task model.

# Framework for Modelling Multiple Input Complex Aggregations for Interactive Installations

Nicolas Padfield
Roskilde University / illutron
HUMTEK, Building 8.1
Universitetsvej 1
Roskilde Denmark
nicolasp@ruc.dk

Troels Andreasen
Roskilde University
CBIT, Building 43.2
Universitetsvej 1
Roskilde Denmark
troels@ruc.dk

***Abstract*—We describe a generalized framework as a method and design tool for creating interactive installations with a demand for exploratory meaning creation, not limited to the design stage, but extending into the stage where the installation meets participants and audience. The proposed solution is based on fuzzy logic and provides a method for variably balancing interaction and user input with the intention of the artist or director. An experimental design is presented, demonstrating an intuitive interface for parametric modelling of a complex aggregation function. The aggregation function unifies hierarchical, importance-weighted and ordered-weighted fuzzy averaging to provide complex combinations of user input.**

***Index Terms*—interactive installations; fuzzy logic; aggregation; parametric modelling; intuitive interface**

## I. INTRODUCTION

My bedroom floor is completely open to a creative process. One can place the socks in any pattern one wishes – but is it art? Or is it just a mess? The Mona Lisa is widely acknowledged as being art. But is the Mona Lisa interactive? Only if you do as vandals did in 1956 (twice), 1974 and 2009 and throw acid, paint or tea mugs at her. Somewhere, between these two endpoints – my bedroom floor and Leonardo da Vincis famous masterpiece – lies the realm of interactive art. An interactive installation is by definition a combination of design choices taken apriori by the artist and input data generated by the audience, users or participants. Usually, the amount of control afforded the participants is predefined. There is a rigid framework, both physical reality and in software, which defines how and to what degree the participants can affect the work of art. But, what if the division between the power of the artist and the power of the participants was not set in stone, but was itself a variable that can be tweaked in real time, can be experimented with and optimised for a particular setting or particular participants?

In this paper, we describe a generalised software framework which allows an artist or designer to tweak the weighting of each individual input and of groups of inputs. We postulate that this provides three advantages over, e.g., programming the system in a traditional way: it enables non-programmer artists and designers to create complex installations, it enables complexity, and enables real time adjustment of the installation itself. By a complex system, we mean a system where it is not easy to predict the exact output even given knowledge of the input and state of the system. While many of the issues described are applicable to a wide range of interactive systems, we will in the following primarily focus on the particular case of an art installation as an example.

Ideally, we wish our installations to attract participants and be easy to start interacting with – at the same time as providing depth and complexity warranting extended exploration. A good explanation of threshold and ceiling and the desirability and difficulty of achieving both low threshold and high ceiling at once can be found in [13] "Threshold and Ceiling: The 'threshold' is how difficult it is to learn how to use the system, and the 'ceiling' is how much can be done using the system. The most successful current systems seem to be either low threshold and low ceiling, or high threshold and high ceiling. However, it remains an important challenge to find ways to achieve the very desirable outcome of systems with both a low threshold and a high ceiling at the same time." [13]

We posit that enabling complexity makes it more likely that participants will engage with the system for a longer time span, wanting to explore the possibilities, as the output is not obviously tightly coupled to the input, and the output may be diverse even given similar input.

The approach to model interactive installations described here aims at providing flexibility for the user/artist and builds on several cases of practical experience with developing installations, among which are [1], [2]. The ideas evolve from earlier concepts such as physically interactive environments [3], [4], immersive virtual environments [5], and "tangible interfaces" [6], originating in Krugers seminal ideas dating back to the 1980s on computer-controlled interactive spaces [7]. Numerous examples appear in the literature of the development of interactive environments for the general purpose of communication, in areas such as advertising, entertainment, story-telling and dissemination of cultural assets; see for example [8], [9], [4]. One of the main motivations for such developments seems to be that physically interactive environments are perceived to offer the user a greater sense of presence and immersion, allowing the user to engage more actively with the content of the communication.

 ISBN: 978-1-61208-235-6 

### A. Why intelligent systems

Nowadays, most audience-interactive art installations are controlled by computers. Thus, behind the scenes, there is always some program or engine which determines the behaviour of the system. Such a system can be relatively simple, with output coupled relatively directly to input or it can be very complex, with output being dependent on multiple factors including current input, past input, predefined data, meta data (data about the data), and even machine learned data derived from the interaction history itself.

Installations of low complexity can be driven by conventional deterministic procedural programs. But, as complexity increases, these dedicated programs quickly become complicated and have the disadvantages of lacking generalisability and not handling uncertainty very well. Installations that choose elements from a large set of data based on myriad user input variables, of which some may be uncertain or conflicting are good candidates for an engine based on some form of artificial intelligence (AI).

### B. Artificial intelligence

While generalised intelligence (strong AI) [10] is still a long term goal for the AI community, applicable AI consists of more specialised systems that are good at a particular task. Artificial intelligence is used for data mining, process control, logistics, diagnosis and in many other areas. The AI field has developed highly successful methodologies for dealing with incomplete or uncertain information, including Bayesian logic [11] enabling probabilistic reasoning in adaptive conditional probability networks, and Artificial neural networks, which are inspired by structural and functional characteristics of biological neural networks, interconnected neurons process information using a connectionist approach.

### C. Fuzzy logic

Fuzzy logic is many-valued logic that generalizes conventional Boolean logic, and enables approximate reasoning [14]. Fuzzy truth values range between the extremes 0 to 1 corresponding to “completely true” and “completely false” and Fuzzy sets elements have a degree of membership, described by a membership function, in the range [0, 1], as opposed to crisp sets where membership is bivalent. Fuzzy logic is useful handling a multitude of sensor inputs, often analogue, sometimes pointing in conflicting directions. The usefulness of fuzzy logic derives from the fact that many problems in the real world, especially ones involving human reasoning, are approximate in nature.

One hundred people drumming is not bivalently either completely in time or out of time or fast or slow. As Lotfi Asker Zadeh wrote in his seminal work *Fuzzy sets* [15]: “More often than not, the classes of objects encountered in the real physical world do not have precisely defined criteria of membership” It is no coincidence that one of fuzzy logics main applications is within industrial process control, that bears more than a passing resemblance to interactive installation process control.

## II. Introducing a generalized fuzzy logic art support framework

In the following, we describe a generalized fuzzy logic framework that facilitates rapid building and experimenting with interactive art, making it feasible to easily create high complexity interactive installations. A system which provides flexible arbitration, balance between interaction and art. Our aim is to enable non technically inclined artists to model complex behavior based on multiple input. Our approach is to apply highly flexible fuzzy aggregation, more specifically hierarchical, importance weighted, ordered weighted aggregation [18], [19], and to provide an “intuitive” user interface that is easy to grasp without fully understanding the mathematical functions behind it.

Interactive art is a balance, of both the artists’ wishes and the participants’ actions. Pre-digitally, the artist could control *what* the participant could do. The artist might have decided that there is a welded steel frame (which it is difficult for the participant to alter), and an inviting brass handle (which is inviting and easy to turn). In the digital age, the artist can also decide and adjust *to what degree* the participant may decide. The participant may turn the handle, but at this moment that weighs in at only 0.3, while the artists’ wishes weigh in at 0.7. Artificial intelligence based on fuzzy logic is a prime candidate for our use because it supports

- Adaptability – suitable for reaching decisions from a number of heterogeneous, possibly conflicting inputs.
- Narrativity – suitable for supporting narrativity, e.g., for searching a large database and selecting *which* media element to become next in a sequence (comprising a story). This is because attributes of elements will often be humanistic and lend themselves to fuzzy quantification better than to binary quantification or procedural programming; and because presuming a finite number of available elements to choose between and multiple, sometimes conflicting inputs, fuzzy is suitable for choosing the best available element.
- Live data – suitable for taking live sensor data and reaching a decision. One can define a parameterized fuzzy linguistic concept such as “rhythmicality”.
- Flexibility in mixing live and predefined data. As fuzzy aggregation can handle both live and predefined data with equal ease, it is possible to use the same aggregations to combine both live and predefined data.
- Realtime adjustment – it is relatively straightforward to tune weighting parameters in real time, giving us ease of experimentation or adjusting the interaction to suit the participants.

### A. Raising the interactivity to another level

An installation programmed in a procedural fashion usually lacks the ability to adjust the interaction control in real time. We would like to let the system run as designed and handle all interaction, while at the same time being able to adjust interaction parameter weightings to experiment and achieve the best interaction scenario for the particular participants and venue. This is a sort of second order version of the Wizard of Oz technique (John F. Kelley, more directly applicably [16]). Ideally not only weights but the very degree of complexity and degree of interactivity should be adjustable parameters. It can be difficult to design perfectly and gauge the audience ahead of time. While the physical characteristics of a work are usually difficult to alter, with suitable algorithms and separation of data and control structures it becomes feasible to adjust the software response in real time. The degree of interactiveness can be a parameter in itself. There are the following possibilities

- 1st. order: the participants input influences the output
- 2nd. order: the participants input alters the installation itself (the machine tunes weighting factors, etc.; machine learning)
- the artists tune weighting factors
- the artists adjust major system parameters such as degree of complexity

## III. Fuzzy sets, fuzzy logic, set operations

Say we want to quantify the enthusiasm of drummers. It is a matter of opinion whether 90 beats per minute is “fast” or “very fast” and it would be counter-intuitive to define 89 bpm as being vastly different from 90 bpm. What is needed here is a smooth transition, from what a human observer would call “fast” to what she would call “very fast”. This can be described more accurately by fuzzy set theory.

Should we wish to determine whether the drumming at a given time or place is fast, we could define “fastness”. We define a fuzzy set “fast” as a subset of the set of possible values for drumming speed as specified by $speed(x)$ for an object (a drummer) $x$ by the following set membership function:

$$fast(x) = \begin{cases} 0 & speed(x) < 80 \\ \frac{speed(x)-80}{50} & 80 \leq speed(x) \leq 130 \\ 1 & speed(x) \geq 130 \end{cases}$$

In reality, membership functions are rarely this simple, and are not always based on a single dimension. It might for example be more intuitive and closer to the intention to define enthusiasm(x), based on both speed(x) and amplitude(x). We show a linear function for simplicity, a sigmoid function might be more suitable for many applications.

In general, a fuzzy set membership function $m_A : X \to [0,1]$ ($X$ being the universe of discourse) has the following properties:

$$m_A(x) \begin{cases} = 0 \text{ if } x \notin A \\ \in \, ]0,1[ \text{ if } x \text{ is partially in } A \\ = 1 \text{ if } x \in A \end{cases}$$

where $A$ is a fuzzy set (for example, elements that meet a particular criteria) and $x$ is a variable $x \in X$ (for example media elements).

Generally, the intersection of fuzzy sets A and B (visualised in Figure 1) is defined by:

$$m_{A \cap B} = \min(m_A, m_B)$$

while the union is:

$$m_{A \cup B} = \max(m_A, m_B)$$

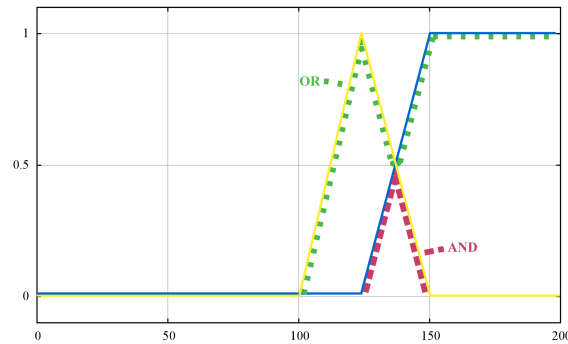



Fig. 1. Red dashed: about 125 AND high. Green dotted: About 125 OR high

The expression $m_A(x)$ can be considered as the truth value of the proposition “x is a member of A” and A can thus be considered as a logic predicate. There is isomorphy between fuzzy set theory and fuzzy logic and for given predicates A and B conjunction A∧B and disjunction A∨B can be defined correspondingly to intersection and union by:

$$m_{A \wedge B}(x) = \min\ (m_A(x), m_B(x))$$

$$m_{A \vee B}(x) = \max\ (m_A(x), m_B(x))$$

### A. Set operations and aggregation

An installation may have inputs that are more or less important, that should be logically grouped, and there may be a limited number of choices (e.g., video clips) where we are interested in the best match given a number of different inputs, which may point in different directions. These requirements can be fulfilled with a layered, grouped approach with a combination of three types of aggregation. We are seeking means to combine inputs in flexible ways and are considering fuzzy set operations for this purpose.

A fuzzy set operation is an operation on fuzzy sets. Fuzzy set operations can be considered a generalization of crisp set operations, many generalizations being possible. The “standard” fuzzy set operations, intersection (conjunction) and union (disjunction), defined above immediately generalizes from 2 to n-argument operations such that for instance n cirteria can be combined by a conjunction:

$$m_{A_1 \vee \ldots \vee A_n}(x) = \max\ _{i=1,\ldots,n}(m_{A_i}(x))$$

These operations are encompassed by more general classes of operations. One important such class is the so called Ordered Weighted Aggregation (OWA). We define OWA below and introduce to generalizations taking importance weighting and hierarchical aggregation into account.

*Ordered Weighted Averaging* is a parameterisable class of mean type aggregation operators first introduced in [18]. The parameters are given as a set of so called order weights that

 ISBN: 978-1-61208-235-6 

apply in the given order to the most, the second most, and so on, fulfilled criteria. An OWA operator is a mapping $F : R_n \rightarrow R$ that has a collection of order weights $W = [w_1, ..., w_n]$ in the range $[0, 1]$ such that:

$$F(a_1, ..., a_n) = \sum_{j=1}^{n} w_j b_j \quad \text{, where } b_j \text{ is the } j^{th} \text{ largest } a_i$$

with $\sum_{i=1}^{n} w_i = 1$ and where $a_i$ is the degree to which the i'th criteria is fulfilled.

OWA enables us to balance artistic intent with media element (for example video clips) scarcity. While strong control of artistic intent might seem to indicate specific control of which criteria is most important is advantageous, in a real world situation with a limited number of media elements, it may mean the total fulfilledness of most criteria is low, because one was weighted as high importance and no media element was available that satisfied both the high importance parameter and the other parameters. With OWA we achieve both a guarantee for all parameters being included and additionally, as we know the sorted order of parameters, we can with weighting factors easily adjust whether highly fulfilled parameters are given most weight, less fulfilled parameters are given most weight, or all parameters are given equal weight.

OWA can be parameterised between $\wedge$ (pure conjunction) and $\vee$ (pure disjunction). The max, arithmetic average, median and min are members of this class. OWA has been widely used in computational intelligence because of the ability to model linguistic expressions.

While order weights relate to the best fulfilled order, *Importance weighting*, on the other hand, relates to specific criteria. Each input, e.g., from a sensor, is multiplied with a weighting factor, enabling us to define that, e.g., one input is twice as important as another input.

*Hierarchical aggregation* is a generalized aggregation introduced in its basic form in [19] that allows the combination of different types of aggregations based on the OWA operator and including importance weighting. Each node in the hierarchy can be attached individual parameters for order and importance weighting. The leaf nodes comprise a grouping of the input and the aggregation at each node delivers input to the parent node. Hierarchical aggregation thus allows us to group inputs, aggregate them in a group and send the result up in a hierarchy to a parent aggregation. This is especially useful when it makes sense to group inputs as there are classes of input that are fundamentally different in nature.

But, where does this leave the artist? There can be something very intuitive about an OWA rather than a logical expression, but how to allow an ordinary user to visualise the possibilities? Ideally we need a knob with complete user control unfettered by artist wishes at one end and complete artist control non-interactivity at the other end.

## IV. Combining our aggregation operators to offer full parametrisation

We are aiming for fully parametric aggregation, adjustable by the artist by intuitive means. Especially, the order weights are difficult to set due to the requirement that they have to sum up to 1. However, we can define a simple function which provides all n order weights based on a single number between 0 and 1 as follows. Having defined such a function, we can introduce a slider in the interface for adjustment allowing an artist to decide on a scale how fulfilled the different criteria must be:

OR <—> AND

with the left extreme corresponding to "any requirements fulfilled" and the right to "all requirements fulfilled". This is an alternative to requiring traditional logical expressions.

Given a function such as:

$$Q(y) = y^{(\frac{1}{p}-1)} \quad , \quad p \in ]0, 1]$$

we can introduce $k$ weights from a single value for p:

$$w_i = Q(i) - Q(i-1) = \left(\frac{i}{k}\right)^{(\frac{1}{p}-1)} - \left(\frac{i-1}{k}\right)^{(\frac{1}{p}-1)}$$

with $i \in \{1, 2...k\}$ and $p \in ]0, 1]$ This is described further in [17]. We want a function, that given one number, gives us $n$ weights out, and they must sum to 1. This method of parametrising a function, calculating all $n$ weights at once using only one input is only useful with OWA. The method is to take the difference between function values, given $x$ values of a multiple of $1/n$ where $n$ is the number of parameters and hence the number of weights required. The function must be monotonically increasing in the interval [0,1] and the sum of the weights must be 1. Any function that fulfils these requirements could potentially be used.

$$w_1 = Q\left(\frac{1}{n}\right) - Q\left(\frac{0}{n}\right) \text{ ... } w_n = Q\left(\frac{n}{n}\right) - Q\left(\frac{n-1}{n}\right)$$

In the concrete case illustrated in Figure 2, for $n = 4$ and $p = 0.3$, the weights are

$$w_1 = 0.039 \quad w_2 = 0.159 \quad w_3 = 0.313 \quad w_4 = 0.489$$

### A. Importance weighted aggregation

Importance weighted aggregation involves giving each parameter a weight, often but not necessarily in the range [0,1]. In an interface, this can be achieved by having a control and display of the importance weight for each parameter - for example sliders or knobs. If all the weights are the same (e.g., 1), the importance weighting is neutral and in effect turned off.

 ISBN: 978-1-61208-235-6 

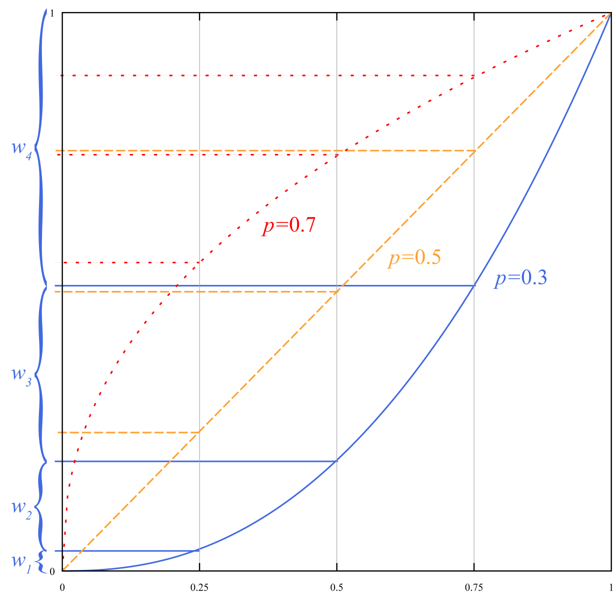



Fig. 2. Finding $n$ OWA weights which sum to 1, for $n = 4$.

### B. Hierarchical

Hierarchical aggregation is a layered, grouped aggregation. In our design it is used as a concrete way of combining grouping, OWA and importance weighted aggregation. A hierarchical aggregation is suitable for grouping. If there are very different parameters (inputs), e.g., sound level and geographical position, a hierarchical aggregation will give a natural grouping. See Figure for an example.

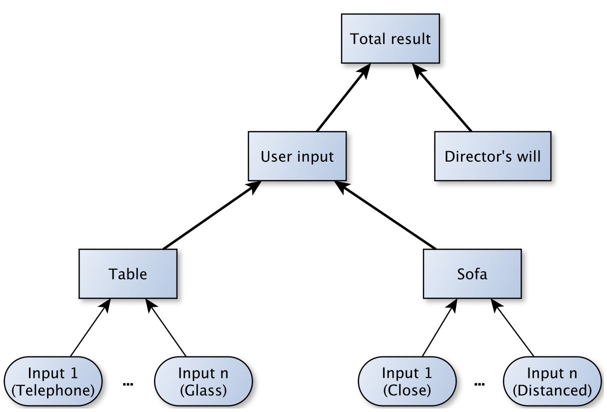



Fig. 3. An example hierarchigal aggregation. Rectangles are aggregations, ovals are inputs.

Here, each rectangle is an aggregation, in our generalised system each is first an importance weighted, then a parametrisable OWA.

### C. Combining

It is quite possible to combine one or many different forms of aggregation, either simultaneously or in nodes of a hierarchy. It is not necessary to provide different aggregation objects as the adjustable parameters provide the possibility to "turn off" any unwanted feature, allowing us to keep it simple, using only one type of aggregation object.

An importance weighted ordered weighted aggregation could look like this: $in$ are inputs, $imp$ are importance weights, then sorting by value, $weight$ are the weighting factors we obtained above. OWA is a method that enables using all inputs instead of just one. This is done by adding all the inputs, but as we want the result of the aggregation to be in $[0, 1]$, we first regulate the inputs by multiplying them by weights. Which input is regulated by which weight is decided by the value of the input, so we sort the inputs before multiplying by the weights. For example, given 4 inputs $a_1, a_2, a_3, a_4$, we sort them by size. The sorted input we call $b_1, b_2, b_3, b_4$, and we then multiply $b_1$ by $weight_1$ so: $(b_1 \times weight_1), (b_2 \times weight_2), (b_3 \times weight_3), (b_4 \times weight_4)$

In this design, the above combined OWA and importance weighted aggregation is used as each node of a user-designable hierarchy, providing a ordered weighted importance weighted hierarchical aggregation. This gives the full freedom to weight parameters (e.g., inputs) by importance, to linearly choose from a range from AND to OR for the OWA aggregation, and to group the parameters in any number of groups and steps. The final output is from the topmost aggregation, the node with no parent.

This fully parameterised design allows the artist to play the system at the meta level (i.e., not altering the inputs, which come from the participants, but more subtly adjusting, in real time, how the inputs are processed. This feature can be augmented by implementing standards based real time adjustment of sliders, etc., enabling use of a physical control surface, e.g., a MIDI (Musical Instrument Digital Interface, a de facto industry standard for sound and interactive installation control) controller.

### D. Machine learning

While the initial goal of this work is to enable the artist/director to easily adjust and experiment with the aggregations of a fuzzy logic based artificial intelligence, the logical next step is to implement adaptive fuzzy logic, machine learning. Our design allows any number of aggregations, sensors or outputs can be routed to any number of inputs, and weights are all in the range [0,1] — all that is missing is to allow an output to control a weight.

While actual implementation is for further work, an overview of required aspects includes a paradigm for combining linguistic and numerical information; choice of learning using, e.g., backpropagation, feedback loops, orthogonal least squares or neighbourhood clustering; and possible inclusion of artificial neural networks in a hybrid neurofuzzy system. Adaptive fuzzy filtering will probably be advantageous given the dynamic and nonlinear nature of many of our systems. All these aspects are active research topics. Adaptive fuzzy logic is often used for systems where it is advantageous to learn a system's characteristics and avoid hysteresis, e.g., maximum power point trackers for solar panel arrays - supplanting, e.g.,

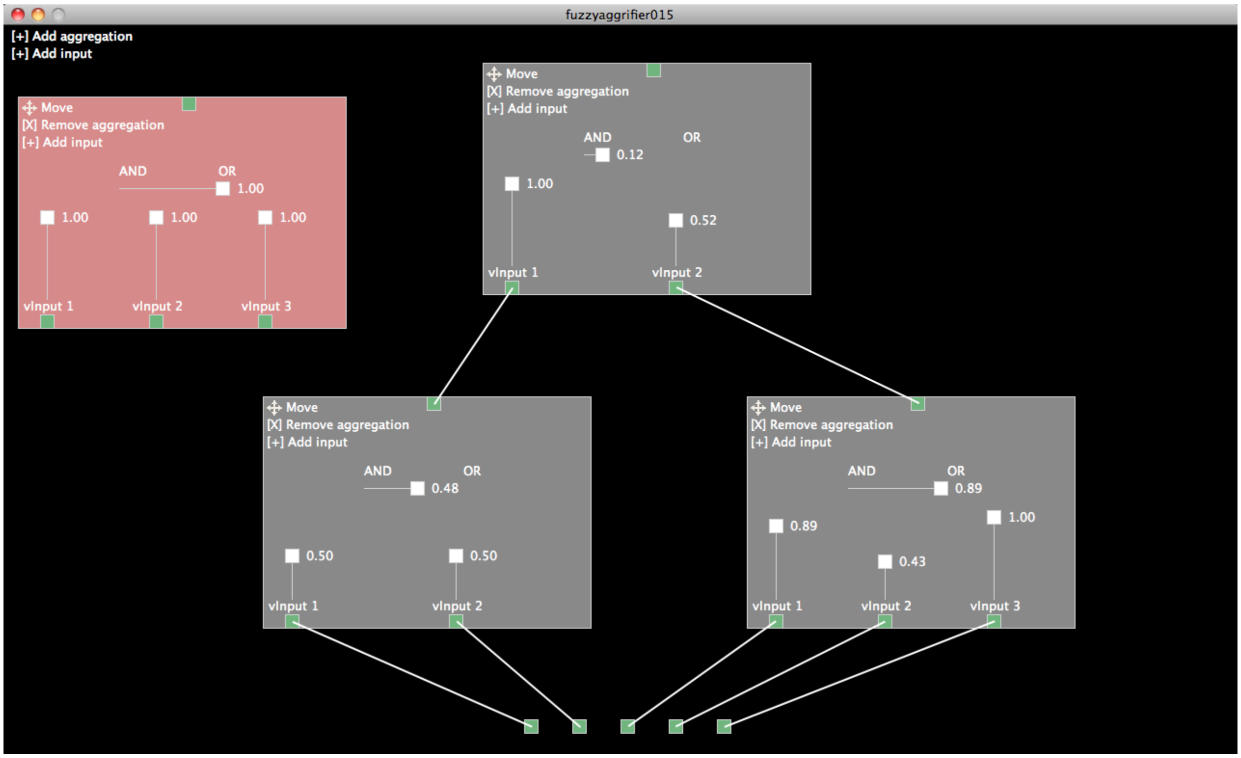



Fig. 4. Experimental interface design iteration 0.8.

PID (proportional-integral-derivative) controllers in diverse applications [12].

## V. Solution: What we want for artists - Interface design rationale and experiments

The requirements for the system are that it:

- Must process the data in real time
- Make it easy to enter data for different projects - separation of data, meta data and control logic is required.
- Handle complexity so great the output is not recognisably deterministic.
- Be second order adjustable in real time (i.e., allow weightings etc. to be adjusted)
- Preferably have easy enough controls (e.g., a hardware control surface) that it invites "playing" and "tweaking" once configured.
- Preferably be understandable by a motivated artist
- Preferably be capable of adjusting itself if so wished

The objective is to design an interface for artists which makes complex fuzzy operations reasonably intuitive. The interface should enable a non computer scientist to design a fuzzy AI system, and tune it in real time. What might seem a daunting task is made easier by the fact that understandable metaphors are available for most steps of the process. Grouping can be symbolised with boxes, hierarchy with lines between these boxes, importance with sliders like an audio mixer or knobs like volume. While a complete system including databases, all kinds of input and output options and machine learning is a major undertaking, the system initial design and programming is progressing well and the interface design has already been through several iterations.

The interface is designed thus: at the bottom are the inputs, any number can be instantiated by pressing "[+] Add input". At the top of each aggregation is an output; if there is no parent this is considered the final output. Any number of boxes=groups=aggregation operations can be instantiated by pressing the "[+] Add aggregation" button. Each box is an object which contains an arbitrary number of inputs, including a slider to define an importance weight for each input. An input can be added by pressing the "[+] Add input" button. It contains methods for an OWA aggregation, a slider for variably parametrically adjusting this aggregation from AND to OR, and one output.

The box is by its nature a grouping. It is often opportune to be able to group inputs. For example, inputs which are in their nature fundamentally different, e.g., microphone volume and geographical position. Lines can be drawn between objects by clicking on a symbolic output jack and then on a symbolic input jack. A sensor can be routed to any number of inputs and the output of an aggregation can be routed to any number of inputs.

Each node in the hierarchy has its own individually controllable OWA and importance weights, just as its place in the hierarchy and what inputs it receives is user controllable. It is possible to mix - a node can perfectly well take input both from sensors and from an aggregation lower in the hierarchy.

Adjusting a slider from AND to OR may be intuitive for some artists and not for others – an investigation of this

 ISBN: 978-1-61208-235-6

is for future work – but it is certainly more user friendly than requiring a non technical user to *a priori* choose from a predefined selection of aggregations, e.g., MAX; MIN; AND; OR; AVG; MEDIAN. Tuning an importance weight is easily understandable, it corresponds to turning the "value" or "volume" of a sensor or input up and down. A hierarchical aggregation is a natural way to support grouped input, and a visual hierarchy is a natural way to visualise the hierarchy and grouping.

### A. Complete solution

The above implemented interface is useful for experimenting with adjusting aggregation; in practical applications, a few more object types are required. Time is a factor, this can be dealt with in three ways: the system can be input driven, request driven or clock/bang driven. Input driven means that when any data arrives, the whole calculation is performed. Request driven would be that another part of the system requests an answer. This could be suitable for, e.g., a video installation, where the system is aware that the present video is nearing its end, and could ask for what video clip to show next. Clock/bang driven would mean that there is a source that propagates a command "do it now" to all relevant objects "Bang" is a term (used in PureData and MaxMSP) for an irregularly timed pulse, e.g., sent only on certain conditions.

The above still leaves us with the aggregations being the result of a snapshot, a moment in time. It would be advantageous to be able to aggregate over time, just as it would be advantageous to be able to let earlier results affect later aggregations - adding feedback and state to the system. Both goals can be accomplished with a multifunctional FIFO (first in first out) object. The FIFO object is a bucket chain, each bucket storing a result. The user can determine the number of buckets and therein the maximum delay. There is an input at the beginning of the FIFO, an output at the end, which provides the input, in order, but delayed by $n$ cycles. Additionally, there is an output from each stage or bucket. It is optional to use these outputs, they are there to enable the aggregation of inputs over time. By making a ten stage FIFO and connecting the ten stage outputs to a ten input OWA aggregation, it is easy to aggregate the last ten sensor inputs over time. Using the FIFO, some "sensors" can be, e.g., metadata about the current media element and some can be state.

## VI. Conclusion and Future Work

We have demonstrated a design for a user friendly interface enabling non technical users to define and adjust in realtime a set of completely parameterizable complex aggregations. We believe an intuitive interface for a hierarchical, importance weighted, ordered weighted aggregation is an original idea. While simple user friendly interfaces offering a flexible alternative to logical expressions (e.g., AND, OR) consisting of one slider in conjunction with a search field [17] have been demonstrated, enabling a user to easily design a complex aggregation and adjust it in real time is to our knowledge new. We have sketched an experimental prototype of the generalized fuzzy logic interface; field evaluation remains for future work.

## References


[1] T. Andreasen, H. Christiansen, J. P. Gallagher, C. Jacquemin, N. Møbius, N. Padfield, S. Siggaard, D. L. Strand, and P. R. Sørensen *Havhingstens Tur til Irland: en interaktiv oplevelsesplatform.* CBIT, Roskilde University, 2011.

[2] T. Andreasen, J. P. Gallagher, N. Møbius, and N. Padfield. *The Experience Cylinder, an immersive interactive platform: The Sea Stallion's voyage: a case study*. In: R. Emonet, A. M. Florea, (ed). AMBIENT 2011, The First International Conference on Ambient Computing, Applications, Services and Technologies. ThinkMind, 2011 pp. 25-31.

[3] J. Montemayor, A. Druin, A. Farber, S. Simms, W. Churaman, and A. D'Amour. *Physical programming: designing tools for children to create physical interactive environments.* CHI. 2002.

[4] C. S. Pinhanez, J. W. Davis, S. S. Intille, M. P. Johnson, A. D. Wilson, A. F. Bobick, and B. Blumberg *Physically interactive story environments.* IBM Systems Journal 39, no. 3 & 4, 2000.

[5] M. Slater and S. Wilbur. *A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments.* Presence 6, no. 6, 1997 pp. 603-616.

[6] J. M. Sales Dias *Natural and tangible human-computer interfaces for augmented environments.* Proceedings of the 26th annual ACM international conference on Design of communication, SIGDOC '08. ACM, 2008, pp. 181-182.

[7] M. W. Krueger *Environmental technology: Making the real world virtual.* Comm. ACM, 1993, pp. 36-37.

[8] M. Danks, M. Goodchild, K. Rodriguez-Echavarria, D. B. Arnold, and R. Griffiths. *Interactive storytelling and gaming environments for museums: The interactive storytelling exhibition project.* Technologies for E-Learning and Digital Entertainment, Second International Conference, Edutainment 2007. Lecture Notes in Computer Science, 2007, pp. 104-115.

[9] D. Grigorovici *Persuasive effects of presence in immersive virtual environments.* Being There: Concepts, effects and measurement of user presence in synthetic environments. Ios Press, 2003.

[10] B. Goertzel and C. Pennachin, eds. *Artificial General Intelligence*, Springer, 2006.

[11] A. Darwiche, *Modeling and Reasoning with Bayesian Networks* Leiden : Cambridge Univ. Press, 2009.

[12] D. Misir, H. A. Malki, and G. Chen, *Design and analysis of a fuzzy proportional-integral-derivative controller*, Fuzzy Sets and Systems. Volume 79, Issue 3, 1996, pp. 297 - 314.

[13] B. Myers, S. E. Hudson, and R. Pausch, *Past, Present, and Future of User Interface Software Tools*, Carnegie Mellon University, ACM Transactions on Computer-Human Interaction (TOCHI) - Special issue on human-computer interaction in the new millennium, Part 1 Volume 7 Issue 1, March 2000. Open access: http://dl.acm.org/citation.cfm?id=344959 [retrieved: 7, 2012]

[14] G. J. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: theory and applications*. Upper Saddle River, NJ: Prentice Hall 1995.

[15] L. A. Zadeh, *Fuzzy Sets*, Information and control (8), pp. 338-353 http://www-bisc.cs.berkeley.edu/Zadeh-1965.pdf [retrieved: 7, 2012]

[16] B. Buxton, *Sketching User Experiences: Getting the Design Right and the Right Design*, Interactive Technologies, Morgan Kaufmann 2007.

[17] T. Andreasen, *Query Aggregation Reflecting Domain-knowledge*, Intelligent systems for information processing - from representation to applications, Bernadette Bouchon-Meunier, Laurend Foulloy, Ronald R. Yager (eds.) Elsevier 2003, pp. 107-116.

[18] R. R. Yager, *On order weighted averaging aggregation operators in multicriteria decision making*, IEEE Transactions on Systems, Man and Cybernetics, 18(1), 1988, pp. 183-190.

[19] R. R. Yager, *A hierarchical document retrieval language* Information Retrieval 3, 2000, pp. 357-377.