Boris de Ruyter Re
David V. Keyson P
Norbert Streitz M
Nikolaos Georganta
Antonio Mana Gon

LNCS 6439

Ambient
Intelligen

First International Joint Co
Malaga, Spain, November 2
Proceedings

Lecture Notes in Computer Science 6439

Boris de Ruyter Reiner Wichert
David V. Keyson Panos Markopoulos
Norbert Streitz Monica Divitini
Nikolaos Georgantas
Antonio Mana Gomez (Eds.)

Ambient Intelligence

First International Joint Conference, AmI 2010
Malaga, Spain, November 10-12, 2010
Proceedings

Springer

Volume Editors

Boris de Ruyter, Philips Research Europe, Eindhoven, The Netherlands
E-mail: boris.de.ruyter@philips.com

Reiner Wichert, Fraunhofer-Institut IGD, Darmstadt, Germany
E-mail: reiner.wichert@igd.fraunhofer.de

David V. Keyson, TU Delft, The Netherlands
E-mail: d.v.keyson@tudelft.nl

Panos Markopoulos, Eindhoven University of Technology, The Netherlands
E-mail: p.markopoulos@tue.nl

Norbert Streitz, Smart Future Initiative, Frankfurt, Germany
E-mail: norbert.streitz@smart-future.net

Monica Divitini, Norwegian University of Science and Technology, Trondheim
E-mail: monica.divitini@idi.ntnu.no

Nikolaos Georgantas, INRIA Paris-Rocquencourt, Le Chesnay, France
E-mail: nikolaos.georgantas@inria.fr

Antonio Mana Gomez, University of Malaga, Spain
E-mail: amg@lcc.uma.es

Library of Congress Control Number: 2010938138

CR Subject Classification (1998): I.2, H.4, H.3, C.2.4, I.2.11, K.4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-642-16916-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-16916-8 Springer Berlin Heidelberg New York

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper 06/3180

Preface

In a world supported by Ambient Intelligence (AmI), various devices embedded in the environment collectively use the distributed information and the intelligence inherent in this interconnected environment. A range of information from sensing and reasoning technologies is used by distributed devices in the environment. The cooperation between natural user interfaces and sensor interfaces covers all of a person's surroundings, resulting in a device environment that behaves intelligently; the term "Ambient Intelligence" has been coined to describe it. In this way, the environment is able to recognize the persons in it, to identify their individual needs, to learn from their behavior, and to act and react in their interest.

Since this vision is influenced by a lot of different concepts in information processing and combines multi-disciplinary fields in electrical engineering, computer science, industrial design, user interfaces, and cognitive sciences, considerable research is needed to provide new models of technological innovation within a multi-dimensional society. Thus the AmI vision relies on the large-scale integration of electronics into the environment, enabling the actors, i.e., people and objects, to interact with their surrounding in a seamless, trustworthy, and natural manner.

For this reason, in 2001 and 2003, a new series of events was established, namely the conference on smart objects (sOc), organized by France Télécom and by the French National Center for Scientific Research (CNRS). In parallel, a second series of events, namely the European Symposia on Ambient Intelligence (EUSAI), was organized in 2003 and 2004 by Philips Research Europe and the University of Eindhoven.

Due to the synergies in this broad community, the organizers of both series agreed to organize the first joined conference of sOc and EUSAI in Grenoble, France, which was called sOc-EUSAI 2005.

The second joint conference of both series was called the European Conference on Ambient Intelligence to reflect its focus on the AmI vision and its application and technological challenges. This AmI-07 conference was held in Darmstadt and expanded the scope of the series by including three different types of contributions: research contributions, case studies and lessons-learned contributions, and industry and socio-economic contributions. Following that line of thinking,"Services for People" was the key theme of AmI-08 in Darmstadt, where "Well-Being and Care" and "Mobility and Logistics" were the two main fields of applications building the setting for technical research contributions, for case studies, for lessons-learned, and socio-economic papers. AmI-09, organized in Salzburg, marked the 10th anniversary of the production of the term Ambient Intelligence. Here the aim was to look back on what had been achieved, review existing solutions, and identify future challenges.

On the other hand, a new track within the AmI series was created with additional landscape presentation and a more visionary view of the domain. Following this visionary approach, this year's conference, AmI-10, brought together the AmI and the AmI.D conferences in the First International Joint Conference on Ambient Intelligence, held in Malaga, Spain.

September 2010

Boris de Ruyter
Reiner Wichert

Message from the Program Chairs

On behalf of the Program Committee for the Research Papers, we would like to welcome you to AmI-10 – the Joint International Conference on Ambient Intelligence. We received 62 full paper submissions, by authors from 24 countries, from which we accepted 24 papers. Additionally we accepted 10 short papers and 5 landscape papers.

Each paper was assigned to three Program Committee members for review. The reviewing phase was followed by a discussion phase among the respective Program Committee members in order to suggest papers for acceptance. The PC Chairs then made the final decision about acceptance based on all reviews, the discussion results and, if necessary, additional reviewing. We would like to thank all members of the Program Committee for their most valuable and highly appreciated contribution to the community by reading submissions, writing reviews, and participating in the discussion phase.

The range of subjects evident in the selected papers covers many important topics in ambient intelligence: from adaptation to the dynamics of the environment, the design of AmI environments, semantics of AmI environments, security and privacy in AmI applications, contextual systems, influencing the users to sensing environments, and architectural issues for AmI systems, which is reflected in the multifarious program of the conference. We hope that you find the papers interesting, informative, and thought provoking and that we will have fruitful discussions, bringing the community a step closer to our AmI vision.

September 2010

Boris de Ruyter
Reiner Wichert

Organization

After three years of the European Conference on Ambient Intelligence, AmI-10 was organized as the First International Joint Conference on Ambient Intelligence, joining the conference series of AmI and AmI.D.

Organization Committee

General Chair

Antonio Maña	University of Malaga, Spain

Organization Chair

Richard Bricaire	Strategies Télécoms et Multimédia, France

Program and Full Papers

Boris de Ruyter	Philips Research Europe, The Netherlands
Reiner Wichert	Fraunhofer IGD, Germany

Short Papers

David V. Keyson	TU Delft, The Netherlands
Panos Markopoulos	TU Eindhoven, The Netherlands

Landscapes

Norbert Streitz	Smart Future Initiative, Germany

Workshops

Monica Divitini	Norwegian University of Science and Technology, Norway
Nikolaos Georgantas	INRIA, France

Posters and Demos

Kristof Van Laerhoven	TU Darmstadt, Germany
Adrian Cheok	Mixed Reality Lab, Singapore

Publicity

Hristo Koshutanski	University of Malaga, Spain

Local Organization

Antonio Muñoz	University of Malaga, Spain

Program Committee

Richard Bricaire	Strategies Télécoms et Multimédia, France
Adrian Cheok	Mixed Reality Lab, Singapore
Monica Divitini	Norwegian Univ. of Science and Technology, Norway
Berry Eggen	TU Eindhoven, The Netherlands
Alois Ferscha	Johannes Kepler Universität Linz, Austria
Nikolaos Georgantas	INRIA, France
Preben Holst Mogensen	Aarhus University, Denmark
Achilles Kameas	Hellenic Open University, Greece
David V. Keyson	TU Delft, The Netherlands
Antonio Krüger	DFKI, Germany
Kristof Van Laerhoven	TU Darmstadt, Germany
Vincenzo Loia	University of Salerno, Italy
Panos Markopoulos	TU Eindhoven, The Netherlands
Antonio Muñoz	Antonio Muñoz, University of Málaga, Spain
Max Mühlhäuser	TU Darmstadt, Germany
Javier Poncela	i2BC, Spain
Boris de Ruyter	Philips Research Europe, The Netherlands
Bernt Schiele	Max-Planck-Institut für Informatik, Germany
Norbert Streitz	Smart Future Initiative, Germany
Reiner Wichert	Fraunhofer IGD, Germany
Woontack Woo	GIST U-VR Lab, South Korea

Reviewing Committee

Dionisis Adamopoulos	University of Piraeus, Greece
Panayiotis Alefragis	TEI of Mesolonghi, Greece
Pablo Anton	Universidad de Málaga, Spain
Juan Carlos Augusto	University of Ulster, UK
Sander Bakkes	Amsterdam Univ. of Applied Sciences, The Netherlands
Alessandro Bastari	Loccioni Group, Italy
Eugen Berlin	TU Darmstadt, Germany
Ulf Blanke	TU Darmstadt, Germany
Marko Borazio	University of Technology Darmstadt, Germany

Christoph Burghardt	University of Rostock, Germany
Jörg Cassens	University of Lübeck, Germany
Andreu Català	Technical University of Catalonia, Spain
Matty Cruijsberg	Delft University of Technology, The Netherlands
Cristina De Castro	IEIIT-CNR, University of Bologna, Italy
Mark de Graaf	Eindhoven University of Technology, The Netherlands
Boris de Ruyter	Philips Research Europe, The Netherlands
Aysegül Dogangün	Fraunhofer IMS, Germany
Jabe Faber	jpf.design, The Netherlands
Babak A. Farshchian	SINTEF, Norway
Florian Floerchinger	SmartFactory KL e.V., Germany
Thomas Frenken	OFFIS, Germany
PILAR Fuster-Parra	University of Balearic Islands, Spain
Damianos Gavalas	University of the Aegean, Greece
Nikolaos Georgantas	INRIA, France
Nuno Guimarães	University of Lisbon, Portugal
Santoso Handri	Nagaoka University of Technology, Japan
Nicolas Hueber	ISL, France
Iker Huerga	CIC tourGUNE, Spain
Bjoern Joop	University of Duisburg-Essen, Germany
Dietrich Kammer	Technische Universität Dresden, Germany
Vassilis-Javed Khan	Breda University of Applied Sciences, The Netherlands
Thomas Kleinberger	Fraunhofer IESE, Germany
Christian Kleine-Cosack	TU Dortmund University, Germany
Philippe Lalanda	Grenoble University, France
Claus Lenz	Technische Universität München, Germany
Vincenzo Loia	University of Salerno, Italy
Cristina Lopez	University Pablo of Olavide, Spain
Artur Lugmayr	Tampere Univ. of Technology, Finland
George Margetis	FORTH, Greece
Panos Markopoulos	Eindhoven University of Technology, The Netherlands
Alexander Meschtscherjakov	University of Salzburg, Austria
Dorothy Monekosso	University of Ulster, UK
Antonio Muñoz	Universidad de Málaga, Spain
Thomas Norgall	Fraunhofer IIS, Germany
Gabriele Oligeri	ISTI-CNR, Italy
Eila Ovaska	VTT Technical Research Centre of Finland, Finland
Were Oyomno	Lappeenranta University of Technology, Finland
Cuong Pham	Newcastle University, UK
Santi Phithakkitnukoon	Massachusetts Institute of Technology, USA
Patricia Rodriguez	ETRA Research and Development SA, Spain
Natalia Romero	Delft University of Technology, The Netherlands
Charalabos Skianis	Univ. of the Aegean, Greece
Norbert Streitz	Smart Future Initiative, Germany

Hiroshi Tanaka	Kanagawa Institute of Technology, Japan
Alfredo Vaccaro	University of Sannio, Italy
Evert van Loenen	Philips Research, The Netherlands
Reinard van Loo	Frequentis AG, Austria
Carlos A. Velasco	Fraunhofer FIT, Germany
Reiner Wichert	Fraunhofer IGD, Germany
Jan Wojdziak	Technische Universität Dresden, Germany
Olga Zlydareva	Eindhoven University of Technology, The Netherlands

Table of Contents

Automating Routine Tasks in AmI Systems by Using Models at Runtime*

Estefanía Serral, Pedro Valderas, and Vicente Pelechano

Centro de Investigación en Métodos de Producción de Software (ProS)
Universidad Politécnica de Valencia, Valencia, Spain
{eserral,pvalderas,pele}@dsic.upv.es

Abstract. One of the most important challenges to be confronted in Ambient Intelligent (AmI) systems is to automate routine tasks on behalf of users. In this work, we confront this challenge presenting a novel approach based on models at runtime. This approach proposes a context-adaptive task model that allows routine tasks to be specified in an understandable way for users, facilitating their participation in the specification. These tasks are described according to context, which is specified in an ontology-based context model. Both the context model and the task model are also used at runtime. The approach provides a software infrastructure capable of automating the routine tasks as they were specified in these models by interpreting them at runtime.

Keywords: task modelling, routine task automation, models at runtime, OSGi.

1 Introduction

Ambient Intelligence (AmI) is a computer paradigm that tries to make real the vision of Mark Weiser [1] where environments are electronically enriched to make them sensitive to user needs. One of the most important goals of building such environments is helping users in their daily lives by automating their routine tasks, also known as behaviour patterns. A behaviour pattern is a set of tasks that are habitually performed when similar contexts arise [2]; hence, context-awareness is essential to executing tasks on behalf of users under the opportune conditions.

Context-aware approaches have made great advances in introducing context into systems. Some of these approaches use this context to automate user actions. They program rules that trigger the sequential execution of actions when a certain context event is produced (e.g. switching on lights when presence is detected). However, these techniques are only appropriate for automating relatively simple tasks [3]. In addition, they generally require a large number of rules that have to be manually programmed [3]. This makes these systems difficult to manage and to understand.

Other approaches have done excellent work in automating user behaviour patterns inferring them from user past actions by using machine-learning algorithms. For instance, if it is detected that: at 8 a.m. the alarm clock goes off, the lights are switched on, the heating of the bathroom is switched on, and, afterwards, the coffee

* This work has been developed with the support of MEC under the project SESAMO TIN2007-62894 and co financed by FEDER, in the grants' program FPU.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 1–10, 2010.

maker makes a coffee; the execution of these tasks will be automatically trigger at 8 a.m. Thus, these approaches adapt to each user by executing its repeated sequences of actions. However, as the above approaches, they neither take into account users' desires (e.g. the repeated execution of an action does not imply that the user wants its automation). Moreover, they have some limitations: (1) the system cannot infer behaviour patterns until it gathers sufficient user past actions; and (2) they can only reproduce the actions that users have frequently executed in the past.

In this work, we tackle the above presented problems. Note that many behaviour patterns are known at design time (such as the above described). Thus, we propose a tool-supported approach that allows these behaviour patterns to be described at design time and automated at runtime when needed. Specifically, we propose a context-adaptive task model where each behaviour pattern is specified with user participation as a hierarchical composition of tasks. These tasks are specified according to context in such a way that they are capable of adapting to it. This context information is described in an ontology-based context model. We use the concept of task not only because it has proved to be effective in user behaviour modelling [4, 5], but also because it is easily understandable to users [6]. This favours the participation of end-users in their behaviour pattern specification, thereby facilitating that their desires and demands are properly taken into account in this specification.

Both the task model and the context model are brought a step further being also used at runtime. The context model is continuously updated according to context changes by using a context monitor. In addition, both models are interpreted by a Model-based behaviour pattern Automation Engine (MAtE) that is in charge of automating the behaviour patterns as specified in the models.

By using this approach, user behaviour patterns can be analyzed in detail before automating them, achieving that tasks can be performed in a more pleasant manner for users and more efficiently in time and energy concerns regardless of whether or not users have performed them before. For instance, using our approach, instead of switching on the bathroom heating at 8 a.m. (like machine-learning algorithms would do), the system could switch it on ten minutes before to reach the optimum temperature when the user takes a shower. In addition, the system could wake him with his preferred music instead of going off the alarm clock, and also check whether it is a sunny day and, if so, raise the bedroom blinds (instead of switching on the light) to save energy. Moreover, the system could wait until the user enters in the kitchen to make coffee so that it was hot when he arrives, as he likes it.

Therefore, the contributions of this work are:

- A novel approach for specifying user behaviour patterns from context-adaptive task descriptions centred in user demands and desires.
- A software infrastructure based on models at runtime for automating the specified user behaviour patterns when needed.

Finally, note that our work is complementary to other approaches such as machine-learning based approaches. On the one hand, our contributions deal with some of their problems, such as the cold-start problem. On the other hand, machine-learning algorithms can be used to extend or modify the behaviour patterns that are automated when users desires or behaviour change in time, as we further discuss in Section 8.

The rest of the paper is organized as follows. Section 2 gives an overview of our approach. Section 3 presents a real case study. Section 4 describes the task model.

Section 5 describes MAtE. Section 6 evaluates our approach. Section 7 presents the related work. Section 8 discusses our further work. Section 9 explains conclusions.

2 An Overview of the Approach

An AmI system provides users with intelligent pervasive services capable of sensing context and controlling the environment devices. We use a model-driven development (MDD) method, which was presented in [7], to automatically generate the Java/OSGi code [8] of these services from a set of models. In this work, we use the functionality provided by these services to automate the behaviour patterns that users want to be automated. To achieve this, our approach consists of the following steps:

1. **Identification of Behaviour Patterns:** Analysts interview users to determine the behaviour patterns that they want to be automated.
2. **Behaviour Pattern Modelling:** First, the analysts specify the identified behaviour patterns by using a context-adaptive task model. As analysts detect context information on which tasks depend, they specify it using the context model presented in [7]. The task model is refined with end-users' participation until they agree with it. Thus, the task model is complemented by the analysts' knowledge, which can contribute to improving the performance of the behaviour patterns; and users' knowledge, which contributes to taking into account their demands and desires. Once the users agree the specified patterns, analysts link them with the pervasive services that can carry out the pattern tasks.

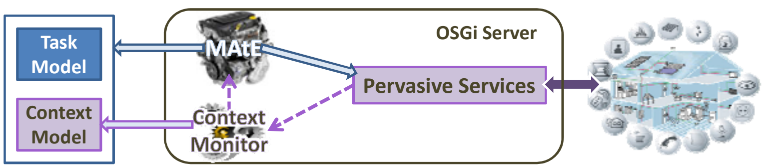


Fig. 1. Global view of our approach at runtime

3. **Running the System** (see Fig. 1)**:** To run the system, an OSGi server [8] is used. The pervasive services, a context monitor and MAtE (the model-based automation engine) are installed and started in this server. The context monitor updates the context model according to context changes and notifies MAtE when a context change is produced. This monitor is out of the scope of this paper; more information about it can be found in [7]. When MAtE receives this notification, MAtE interprets the context model and the task model at runtime to check whether there is any behaviour pattern that has to be executed, and, if so, executing the corresponding pervasive services as specified in the task model.

3 Identification of Behaviour Patterns: A Motivation Example

The philosophy of our approach can be applied to any domain where routine tasks are carried out. In this work, we apply our approach to the smart home domain. A married couple (we will refer to them as Bob and Sarah), was first interviewed to determine their domestic behaviour patterns. We next present two of the most representative behaviour patterns of the 14 identified for the case study:

1. **WakingUp:** At 7:50 a.m. on working days, the system switches on the bathroom heating. Ten minutes later, the system wakes Bob up with his preferred music. Then, if it is a sunny day, the bedroom blinds are raised; otherwise, the bedroom light is switched on. Afterwards, when Bob enters in the kitchen, the system makes a coffee.
2. **GoingOut:** When Bob and Sarah left the house, the system closes all the windows and doors, and switches off all the lights. If it is a working day, the system puts the air and heating conditioner in energy-saving mode; otherwise, the system switches it off, turns the water off at the mains and turns the gas off. Finally, the system reminds users to enable security (they prefer to activate the security themselves; therefore, the system only sends them a notification).

4 Modelling of User Behaviour Patterns

In order to specify the behaviour patterns that users want to be automated we propose a context-adaptive task model. This model is based on the *Hierarchical Task Analysis (HTA)* technique [9], which breaks down tasks hierarchically into other tasks. We propose defining a task hierarchy for each behaviour pattern. The root task of the hierarchy represents the behaviour pattern and has associated a context situation, which defines the context conditions whose fulfilment enables the execution of the behaviour pattern. This root task can be broken down into *Composite Tasks (*which are intermediate tasks*)* and/or *System Tasks* (which are leaf tasks). The Composite Tasks are used for grouping subtasks that share a common behaviour or goal. The System Tasks have to be associated to pervasive services that can perform them. For instance, if a system task to turn on the radio has been defined, this task is associated to a pervasive service that executes this action by interacting with the radio.

In addition, Composite and System tasks can have a context precondition, which defines the context situation that must be fulfilled so that a task is performed (if the precondition is not fulfilled, the task will not be executed). Tasks inherit the context preconditions of their parent task. Also, each task is defined by a task name (which is a sentence that explains the function of the task in a user compressible way) and an internal Task ID (which is a unique identifier). As examples, Fig. 2 shows the modelling of the *WakingUp* and *GoingOut* patterns (see Section 3).

To break down a behaviour pattern or a composite task into simpler tasks we propose two task refinements: the exclusive refinement (represented by a solid line) and the temporal refinement (represented by a broken line). The exclusive refinement decomposes a task into a set of subtasks, in such a way that only one subtask will be executed (disabling the others). For instance, in the *GoingOut* pattern, the *manage home resources* task is refined in two tasks by exclusive refinements in such a way that only one of these tasks will be executed (depending on whether it is holiday or not). The temporal refinement also decomposes a task into a set of subtasks, but this refinement provides constraints for ordering the execution of these subtasks. Using this refinement, the subtasks of the same parent can be linked using temporal operators. For reasons of brevity, we only present the operators that are shown in Fig. 2. To define them, we based on CTT [4], which provides one of the richest sets of temporal operators:

- T1>>T2, enabling: the *T2* task is triggered when the *T1* task finishes; e.g., the task of the *GoingOut* pattern for notifying users to activate security is only triggered when the tasks for managing the home resources have been performed.
- T1|||T2, concurrent tasks: T1 and T2 can be performed in any order; e.g., the execution order of the *turn on the radio* and *lighting* tasks of the *WakingUp* pattern is not relevant.

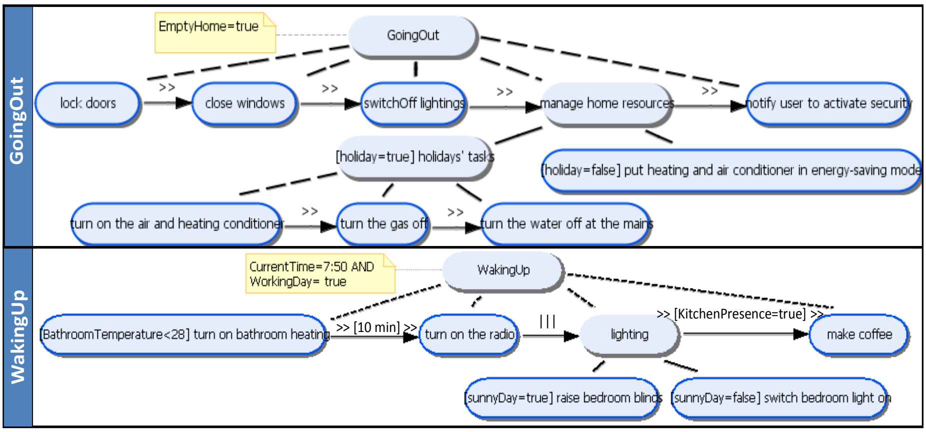


Fig. 2. Examples of behaviour pattern modelling (graphical representation)

In addition, to properly capture the pervasive system automation requirements, we add the following operator:

- T1>>[c]>>T2, enabling when the *c* is fulfilled: after executing T1, T2 is enabled when the condition *c* is fulfilled. *c* can be a temporal restriction (e.g. 3 minutes after the *wind up awnings* task finishes, the *switch sprinklers off* is enabled) or a context condition (e.g. after the sprinklers have been switched off, when it stops raining, the *modify irrigation system* task is enabled).

```
<?xml version="1.0" encoding="UTF-8"?>
<org.pros:TaskModel xmi:version="2.0" xmlns:xmi="http://www.omg.org/XMI"
…
<Task xsi:type="org.pros:SystemTask" name="turn on bathroom heating" ID="WakingUp_TOBAHC"
isChildOf="//@Task.0">
       <ContextPrecondition ContextPreconditionString="BathroomTemperature<28">
       <TemporalRelationship TemporalRelationshipType=">>[10 min]>>"
       TemporalRelationshipTo="//@Task.2" TemporalRelationshipFrom="//@Task.1"/>
    <service href="SmartHomeServices.serviceModel#//@Service.3"//@Method.switchOnHeating"/>
</Task>
<Task xsi:type="org.pros:SystemTask" name="turn on the radio" ID="WakingUp_TOR"
isChildOf="//@Task.0">
    <TemporalRelationship TemporalRelationshipType="|||"
       TemporalRelationshipTo="//@Task.3" TemporalRelationshipFrom="//@Task.2"/>
    <service href="SmartHomeServices.serviceModel#//@Service.2"//@Method.turnOnRadio"/>
</Task>
…
</org.pros:TaskModel>
```

Fig. 3. Part of the XMI representation of the *WakingUp* behaviour pattern

To specify the context conditions (in context situations that enables the execution of the pattern, in task preconditions and in relationships), we use a logical expression. This expression combines any number of basic expressions linked by the following logical connectives: and (AND), or (OR), not (NOT), equalities (=), inequalities (!=) and more than (>), or less than (<). The variables used in these expressions have to be previously specified in the context model.

Task Model Implementation: The task model is specified by means of a modelling tool developed using the Eclipse platform and the EMF and GMF plugins [10]. This tool allows the model be graphically edited (as shown in Fig. 2), and also stored in XMI (XML Metadata Interchange). To be able to use this model at runtime, we use its representation in XMI. Fig. 3 shows part of the XMI representation of the *WakingUp* pattern, where the properties of the *turn on bathroom heating* and *turn on the radio* tasks are shown. Note that the *service* property of each system task is a reference to the service in charge of executing it.

5 Automating User Behaviour Patterns: MAtE

In order to automate the specified behaviour patterns when needed, we have developed a Model-based behaviour pattern Automation Engine (MAtE). MAtE provides an API that implements the model management operations needed for interpreting both the context model, to check a context condition, and the designed task model, to execute the pattern tasks according to their specification. Specifically, this API provides a Java class for each element that can appear in a task model (such as *BehaviourPattern, SystemTask, Condition*, etc.). Each class implements methods for obtaining any information about the element that the class represents. To implement these classes we have used the EMF and EMF Model Query (EMFMQ) plugins of Eclipse [10].

In addition, some of the API Java classes represent context conditions, such as *ContextSituation* or *ContextPrecondition*. These classes provide the *checkCondition* method as well, which queries the context model to check the context condition. To do this, the method interprets the logical expression of the condition and builds a query by using SPARQL [11] (which is a graph-matching query language for ontologies recommended by the W3C). Once the query has been built, the method uses the Pellet reasoner [12] to launch it against the context model.

Using this API, MAtE automates the specified behaviour patterns by following two steps, which are shown in Fig. 4:

1. The first step is to **check the fulfilment of the context situations** specified in the task model. The context monitor [7] (which is in charge of updating the context model according to context changes) notifies MAtE when a context change is produced. MAtE then gets the context situations that depend on that change and then checks whether they are satisfied. To do this, MAtE uses the *checkCondition* method of the *ContextSituation* API class. For instance, every minute, the context monitor updates the *CurrentTime* context property and notifies MAtE. Since the context situation of the *WakingUp* pattern depends on this property, MAtE gets this context situation and checks whether it is then satisfied.
2. The second step is executed **when a context situation is satisfied**. When this happens, **the behaviour pattern related to that context situation is first obtained** by using the *getBehaviourPatternByContextSituation* method of the

TaskModel API class, **and then executed** by using an algorithm that MAtE implements. This algorithm executes the pervasive services related to the system tasks of the corresponding pattern according to its refinements, the temporal relationships specified among its tasks and the up-to-date context information (stored in the context model) on which tasks and relationships depend. To execute each system task, its related pervasive service (which is running in the OSGi server) is searched by using the OSGI capabilities [8], and then executed. For instance, if the *WakingUp* context situation is satisfied (see Fig. 2), the algorithm checks if the bathroom temperature is less than 28 degrees, which is the context precondition of the first system task. If so, its related pervasive service, i.e., the *switchOnHeating* method of the *BathroomHeating* service (see Fig. 3), is searched and executed. Ten minutes later, the algorithm executes the pervasive service related to the next system task, i.e., the *turnOnRadio* method of the *Radio* service. Then, the algorithm checks if it is a sunny day. If so, the *raiseBlinds* method of the *Blind* service is executed; otherwise, the *switchLightsOn* method of the *BedroomLighting* service is executed. Finally, when the user enters in the kitchen, the *makeCoffee* method of the *CoffeeMaker* service is executed.

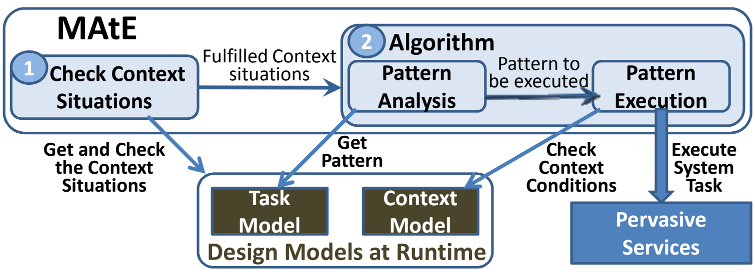


Fig. 4. MAtE process overview

6 Evaluation

We performed a case study-based evaluation to test our approach in supporting behaviour patterns' automation. We developed three case studies to support the automation of behaviour patterns in a smart home. As an example, we presented one of these case studies in Section 3. To develop them, we first interview users to identify the behaviour patterns to be automated. Next, we specified these patterns using the context model and the task model. After teaching the users about the main concepts of the task model, we refine the models until they agreed with the specified patterns. We found that this model was intuitive enough for the users and expressive enough to model the identified behaviour patterns. However, some users found it a little difficult to understand some temporal operators, like >>[c]>>.

To evaluate the feasibility of our approach, we ran the system in real-life deployment sessions. In the experimental set-up, a scale environment with real devices [13] was used to represent the smart home. This execution environment was made up of the OSGi *Prosyst Embedded Server 5.2.* [14] running in a Pentium 4, a network of EIB devices [15], and a device simulator. The EIB network provided users with all the real devices that users could interact with. This network was connected to the PC by an USB port. The pervasive system accessed this network by means of the EIB bundle provided by Prosyst. The Device Simulator, which was presented in [13], simulates by software the behaviour of the rest of devices needed for the case studies.

To deploy the system, we packaged the pervasive services, MAtE, the context monitor and the simulator into bundles, and we installed and started them in Prosyst. To prove that the system automated the behaviour patterns according to their specification in the task model, we implemented a set of tests in which each combination of context values that influences any specified context situation is simulated to see how the system reacted. The user behaviour patterns were correctly automated according to their specification. However, we realized that when more than one pattern had to be executed at the same time (their context situations are fulfilled), the users may want these patterns to be executed in certain order. Therefore, we are currently studying the ways to extend our approach to support the definition of temporal relationships between patterns.

7 Related Work

In this work, we have presented an approach for automating user behaviour patterns from their specification by using context-adaptive task descriptions. Previous approaches to deal with this challenge can be grouped into: machine-learning approaches, rule-based context-aware systems, and task-oriented computing systems.

Machine-learning approaches infer user behaviour patterns from historical data and then automate them. Some outstanding examples are: the MavHome project [16] or the iDorm project [17]. These approaches present some limitations: 1) they require a great amount of training data (the cold-start problem); 2) lack of knowledge about user performed tasks and user desires may lead to automating tasks for which the users may not want automation; and 3) they can only reproduce the actions that users have executed in the past. In this work, we tackle these problems. We provide the system with a set of tasks that are automated since it starts to run. These tasks are described by analysts with user participation, in such a way that: only those tasks that users do want are automated; users' desires and time and energy concerns can be taken into account; and tasks can be automated regardless of whether or not the users performed them in the past.

Rule-based context-aware approaches programme rules to automate user actions when certain context condition arises. Some examples are the works proposed in [18] and [19]. Both works allow users to configure the specified rules by modifying their preferences; however, these approaches do not favour user participation in the design of the rules. Also, these techniques are not appropriate for automating user complex tasks and generally require large numbers of rules which have to be manually programmed [3]. In contrast, in our approach, all the routine tasks that the system automates are described and managed by using the task model. This model facilitates the participation of users in its design by using concepts close to them [6].

Task-oriented computing systems use task modelling to facilitate the interaction of users with the system, since the concept of task is more understandable for end-users [6]. These systems have proved that task modelling is effective in several fields such as user interface modelling [4], assisting end-users in the execution of tasks [20, 21], etc. These works show the growing usage of task modelling and its remarkable results and possibilities to model system behaviour. However, none of these works attempt to automate user behaviour patterns. Therefore, the proposed task models do not provide either enough expressivity (such as relationships between tasks, or context situations) for modelling behaviour patterns, or support for their automatic execution at runtime.

8 Discussion of Our Current and Further Work

This work supports the automation of behaviour patterns that can be known at design time; however, user behaviour may change after a specific period of time and the defined behaviour patterns may become useless for users. To avoid this, these patterns have to be evolved to adapt to user needs. In our approach, the models are directly interpreted at runtime to execute the corresponding patterns as specified; therefore, as soon as the models are modified, the changes are applied by the system. This provides two immediate benefits to carry out the specified pattern evolution: 1) we do not have to maintain the consistency between the system modelling and its implementation as modifications are applied; 2) the evolution can be managed at a high level of abstraction (modelling level) focusing away from lines of code. Taking advantage of these benefits, we are extending the API implemented by MAtE to provide mechanisms for managing the task model at runtime.

Developers will be able to directly use this API in order to adapt the models. However, we think that this adaptation should be performed either in an automatic way or by the end-users. Thus, we are also extending our approach by:

- End-user tools: these tools can provide end-users with user-friendly interfaces to easily change the behaviour patterns that are automated. These interfaces can use the extended API to update the models according to the changes.
- Support to use machine-learning algorithms: when the system is running, the context monitor stores the user actions in the context model. Machine-learning algorithms can use this information to detect changes in the specified behaviour patterns. To achieve this, we are developing a Java interface to properly get this information from the context model. Also, we are studying the development of an end-user tool that periodically shows to users the changes detected by the algorithms and allow users to apply these changes if they so consider. To do this, the tool will use the extended API.

In addition, if the behaviour pattern evolution requires new services to carry out the new tasks or to sense context not considered yet, they can be easily developed by using the MDD method [7], and installed in the OSGi Server at runtime [8].

Furthermore, our approach can be applied in very varied environments [8], because our software infrastructure is implemented by using Java and OSGi. We are currently developing a case study for automating and improving the irrigation and fertilization in an orange field. Some of the routine tasks that have been identified to be automated are the following: *periodicIrrigation* and *periodicFertilization* to periodically irrigate and fertilize the land according to its humidity and the season; and *frostSecurity,* to irrigate the land when a frost has been predicted to avoid that the oranges freeze.

9 Conclusions

In this work, we have presented and evaluated a novel approach for confronting the challenge of automating user behaviour patterns. To do this, we have presented an approach based on models at runtime. In this approach, we have proposed a context-adaptive task model that provides a rich expressivity to specify behaviour patterns and is intuitive enough to be understood generally by users. We have also designed and implemented a software infrastructure to support the automation of these behaviour patterns by interpreting the models at runtime. In addition, we are extending this infrastructure to allow the behaviour patterns specified in the models to be evolved.

We believe that using the models in which the behaviour patterns are designed also at runtime can provide us with a rich semantic base for runtime decision-making and with great benefits to evolve the patterns.

References

[1] Weiser, M.: The Computer of the 21st Century. Scientific American 265 (1991)
[2] Neal, D.T., Wood, W.: Automaticity in Situ: The Nature of Habit in Daily Life. In: Psychology of action: Mechanisms of human action (2007)
[3] Cook, D.J., Das, S.K.: Smart environments: technologies, protocols, and applications (2005)
[4] Paternò, F.: ConcurTaskTrees: An Engineered Approach to Model-based Design of Interactive Systems, pp. 483–500
[5] Masuoka, R., Parsia, B., Labrou, Y.: Task Computing-The Semantic Web Meets Pervasive Computing. In: Fensel, D., Sycara, K., Mylopoulos, J. (eds.) ISWC 2003. LNCS, vol. 2870, pp. 866–881. Springer, Heidelberg (2003)
[6] Johnson, P.: Tasks and situations: considerations for models and design principles in human computer interaction. In: HCI International, pp. 1199–1204 (1999)
[7] Serral, E., Valderas, P., Pelechano, V.: Towards the Model Driven Development of context-aware pervasive systems. In: Pervasive and Mobile Computing (2010)
[8] OSGI, http://www.osgi.org/
[9] Shepherd, A.: Hierarchical Task Analysis. Taylor & Francis, London (2001)
[10] Eclipse Platform, http://www.eclipse.org
[11] SPARQL Query Language (2008), http://www.w3.org/TR/rdf-sparql-query/
[12] Sirin, E., Parsia, B., Grau, B.C., et al.: Pellet: A practical OWL-DL reasoner. JWS (2007)
[13] http://oomethod.dsic.upv.es/labs/projects/pervml
[14] Prosyst, http://www.prosyst.com/
[15] EIB, http://www.knx.org/
[16] Youngblood, G.M., Cook, D.J., Holder, L.B.: Managing Adaptive Versatile Environments. In: Pervasive and Mobile Computing (2005)
[17] Hagras, H., Callaghan, V., Colley, M., et al.: Creating an Ambient-Intelligence Environment Using Embedded Agents. IEEE Intelligent Systems 19(6) (2004)
[18] García-Herranz, M., Haya, P.A., Esquivel, A., et al.: Easing the Smart Home: Semi-automatic Adaptation in Perceptive Environments. JUCS 14 (2008)
[19] Henricksen, K., Indulska, J., Rakotonirainy, A.: Using context and preferences to implement self-adapting pervasive computing applications, SPE (2006)
[20] Huang, R., Cao, Q., Zhou, J., et al.: Context-Aware Active Task Discovery for Pervasive Computing. In: ICCSSE (2008)
[21] Sousa, J., Poladian, V., Garlan, D., et al.: Task-based Adaptation for Ubiquitous Computing. IEEE Transactions on Systems, Man, and Cybernetics 36, 3 (2006)

Service Obtrusiveness Adaptation*

Miriam Gil, Pau Giner, and Vicente Pelechano

Centro de Investigación en Métodos de Producción de Software
Universidad Politécnica de Valencia
Camino de Vera s/n, 46022 Valencia, Spain
{mgil,pginer,pele}@pros.upv.es

Abstract. Increasingly, mobile devices play a key role in the communication between users and the services embedded in their environment. All these services compete for the attentional resources of the user. Thus, it is essential to consider the degree in which each service intrudes the user mind (i.e., the obtrusiveness level) when services are designed. In this work we introduce a method for the development of mobile services that can be adapted in terms of obtrusiveness. That is, services can be developed to provide their functionality at different obtrusiveness levels by minimizing the duplication of efforts. In order to define the obtrusiveness level adaptation in a declarative manner we make use of Feature Modeling techniques. An experiment was conducted in order to put in practice the proposal and evaluate the user acceptance for the way in which services are presented.

Keywords: Obtrusiveness adaptation, feature modeling, interaction adaptation.

1 Introduction

The pervasive computing paradigm envisions an environment full of embedded services. With the inclusion of pervasive technologies such as sensors or GPS receivers, mobile devices turn into an effective communication tool between users and the services embedded in their environment (e.g. users can compare prices of products directly interacting with the mobile device). Since mobile devices provide a rich contextual information about the user, the system can anticipate some of the user tasks. However, a complete automation is not always desirable [15]. For certain tasks, some users prefer that the system acts silently in order not to be disturbed. For other tasks, users want to know what is happening behind the scenes. For example, when the favorite program of a certain user begins, the system should consider whether to start recording and/or informing the user depending on the context. If the system decides to inform the user first, it must choose the most adequate mechanism from all the ones available in his/her mobile device (sound, vibration, a text message, etc.).

* This work has been developed with the support of MICINN under the project SESAMO TIN2007-62894 and co-financed with ERDF, in the grants program FPU.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 11–20, 2010.

Since user attention is a valuable but limited resource, an environment full of embedded services must behave in a considerate manner [7], demanding user attention only when it is actually required. Evaluating Presto [8] (a context-aware mobile platform that allows to support different workflows by interacting with the physical environment) we highlighted the need for mechanisms that adapt the degree to which interaction intrudes on user attention.

Towards creating interfaces that adapt their level of intrusiveness to the context of use, several initiatives have studied strategies to minimize the burden on interruptions [9]. However, these initiatives are almost exclusively focused on evaluating the adequate timing for interruptions , while user interface adaptation has received few attention. Other initiatives that focus their works on development context-sensitive user interfaces ([3], [2]) do not consider user attention in the adaptation process.

The main contribution of this work is a method for the development of mobile services that can be adapted to regulate the service obtrusiveness (i.e., the extent to which each service intrudes the user's mind). In this work, we provide a technique that allows to define pervasive services that can be presented at different obtrusiveness level without duplicating efforts in the development. Feature modeling techniques [5] are applied to describe the commonalities and differences between the interaction mechanisms provided for each service in a declarative manner. In this way, the interaction could be adapted to the user attention without explicitly defining each combination of context.

The remainder of the paper is structured as follows. Section 2 introduces the techniques applied in our proposal. Section 3 describes the application of the proposal in a case study that is based on a Smart Home scenario and presents the results from this application. Section 4 presents related work. Finally, Section 5 concludes the paper.

2 Adaptation for Pervasive Service Interaction

This section introduces a technique to adapt the way in which a pervasive service is accessed by adjusting its obtrusiveness level. For defining such services we need to (1) determine the obtrusiveness level required for the interaction and (2) make use of the adequate interaction mechanisms to provide the functionality according to the obtrusiveness level.

In the following sections we provide a detailed description of the obtrusiveness concept and Feature models as they have been defined for the present work, and then we present the technique for adapting the interaction accordingly.

2.1 Obtrusiveness Concept

We make use of the conceptual framework presented in [11] to determine the **obtrusiveness level** for each interaction in the system. This framework defines two dimensions to characterize implicit interactions: *initiative* and *attention*. According to the *initiative* factor, interaction can be *reactive* (the user initiates the interaction) or *proactive* (the system takes the initiative). With regard to the attention factor, an interaction can take place at the *foreground* (the user is fully conscious of the interaction) or at the *background* of user attention (the user is unaware of the interaction with the system). In this work we assume that the events that are generated by a pervasive system can be situated in a different position of the obtrusiveness space depending on the context.

For the application of our proposal, we introduce an order in the values that define the *initiative* and *attention* axes. On the one hand, the extreme values for the *attention* axis are *Background* and *Foreground*. Since this axis represents user attention demands, we could order these values as $Background < Foreground$ to indicate that *Foreground* interactions require more attention than *Background* interactions. On the other hand, the *initiative* axis is related to automation, so we consider that the *Reactive* value provides a lower degree of automation than the *Proactive* value (i.e., $Reactive < Proactive$).

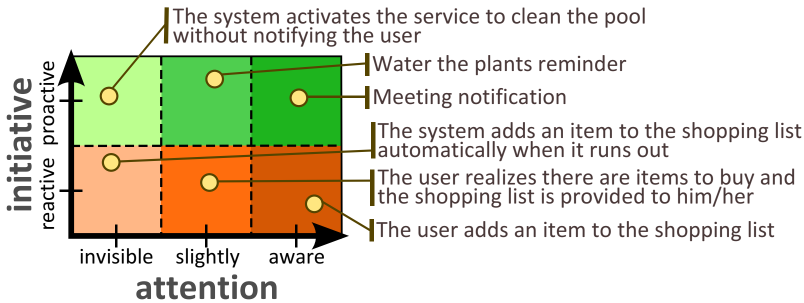


Fig. 1. Events at different obtrusiveness level

Figure 1 illustrates an example of different events in the obtrusiveness space. The initiative axis in this case is divided in two parts: *Reactive* and *proactive*. The attention axis is divided in three segments which are associated with the following values: *Invisible* (there is no way for the user to perceive the interaction), *slightly-appreciable* (usually the user would not perceive it unless he/she makes some effort), and *user-awareness* (the user becomes aware of the interaction even if he/she is performing other tasks). Designers can divide the obtrusiveness space into many disjoint fragments as they need to provide specific semantics.

In the example, we have considered that it is important to inform users about the events that require some actions for them to be performed or events that are urgent. This is why events such as meeting notification or water the plants reminder are represented in a proactive manner (the system takes the initiative). However other events that are more common (adding an item to the shopping list or showing the items of the shopping list) are not announced to the user (i.e., they are marked as reactive in terms of attention). The example of Fig. 1 defines each service at a different attention level according to different context factors such as the message urgency, the location of the user and the location of the mobile. For example, the system is more likely to interrupt the user to notify a meeting (speech interaction can be used) than it is to notify that plants should be watered (a subtle signal can be used).

2.2 Feature Modeling

Feature Modeling is a technique to specify the variants of a system in terms of features (coarse-grained system functionality). The relevant aspects of each platform and the possibilities for their combinations are captured by means of the feature model. Features are hierarchically linked in a tree-like structure through variability relationships such as optional, mandatory, single-choice, and multiple-choice.

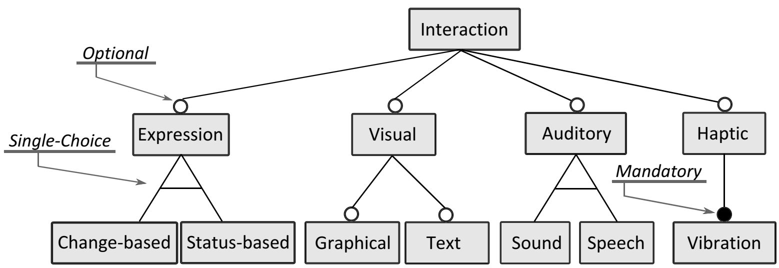


Fig. 2. Interaction mechanisms Feature Model

Besides describing the relevant aspects to the system, feature models have proven to be effective in hiding much of the complexity in the definition of the adaptation space [4]. We make use of Feature Models to describe the possible interaction mechanisms and the constraints that exist for their selection. For example, according to our model showed in Fig. 2 an *auditory* element must either be *speech* or *sound*. In the same way, information or feedback can either be expressed *change-based* (it reports only the changes) or *status-based* (it is continually informing about the status).

The Feature Model of the Fig. 2 is only an example, a designer should describe the possibilities for communication according to abstract aspects that can be mapped into different concrete representations depending on the device used.

Feature models allow us to decompose the interaction in different aspects without explicitly having to define it for each possible combination of context conditions. This avoid duplicating efforts in the development.

2.3 Mapping to a Concrete Interface

In order to choose the most adequate interaction mechanisms according to the obtrusiveness level in each context, we have defined a technique of several steps. The feature concept has been introduced previously to reflect the terms in which the interaction is perceived in an abstract way. These features will be mapped to a set of *"widgets"* and interaction mechanisms that will make up the final UI. In order to allow a flexible composition of the UI elements, we assume a user interface model organized in a *tree structure*. In this structure, components can be contained in other components following a hierarchical representation that allows an easier definition of UIs and an easier support for animation, multi-touch interactions and visual effects as seen in iPhone or Android UIs. This node-based user interface provides an easier node substitution (to adapt UIs at run-time) and an advanced management of interaction events.

In our work, the nodes represent concrete interaction objects. They are any UI components that the user can perceive such as graphical objects, text, image viewers, UI controls, video viewers, etc.

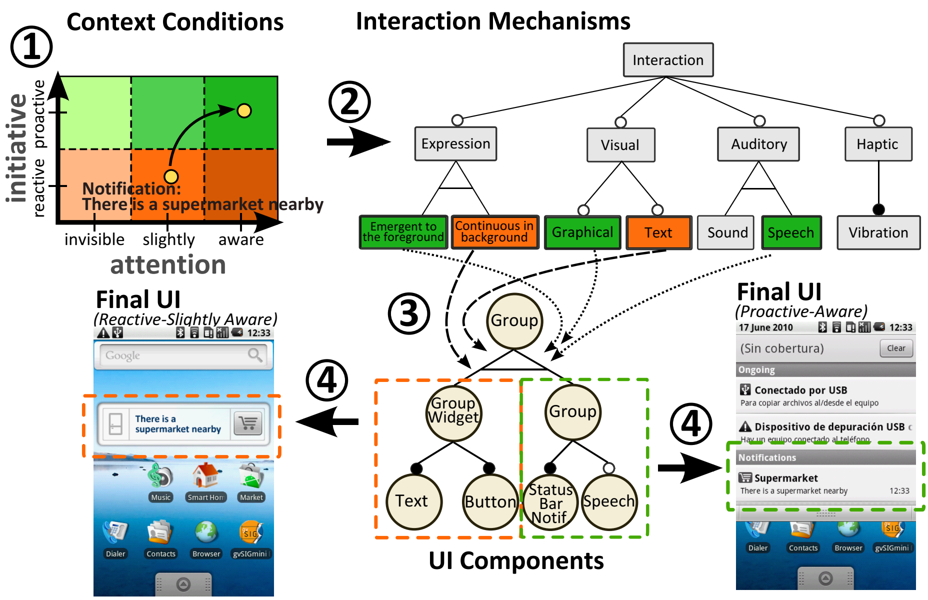


Fig. 3. Approach overview

The overall adaptation steps are outlined in Fig. 3 and described below.

1. **Definition of the obtrusiveness level.** First, the designer determines the obtrusiveness level required for each event depending on the context conditions. Nevertheless, this level can evolve due to changes in the context. For example, the obtrusiveness level for the notification of a supermarket nearby can be changed depending on the user's location, but can be also changed depending on the priority it has for the user (e.g., demanding more attention when the supermarket is closer or when items to buy exceed a fixed number).
2. **Mapping to a Feature Model.** Designers must define the interaction of each obtrusiveness level. This is done through the mapping between each fragment in the obtrusiveness space into a set of interaction features representing interaction mechanisms available. These set of features are interaction aspects preferred for a specific obtrusiveness level. For example, when an event is in the *aware-proactive* space, interaction is offered in a *graphical*

and *speech* manner and the feedback is *change-based* which means that only the changes are reported.

3. **Mapping to the concrete elements.** Then, designers must define how each feature is specified in the concrete interface model. To achieve this, each feature is mapped into a set of nodes representing concrete interaction objects. This determines which UI components must be used to support each feature.
4. **UI composition.** The final UI is made up by fragments of UI whose features are activated from the node-based structure.

Figure 3 shows an example of a mapping from a region in the obtrusiveness space to a concrete node-based interface. In the obtrusiveness space of context conditions we can see an event evolving from a reactive space to a proactive one (e.g., caused by a change in the user's location). In this particular example, the notification of a supermarket nearby was in a *reactive-slightly* space because the user was at a distance too far from the supermarket. For this region in the obtrusiveness space, subtle interaction is preferred, activating the elements *status-based* and *text* and producing the final UI, showed at the left of the Fig. 3. As the user moved closer to the supermarket, this event could evolve requiring other interaction methods more explicit such as *change-based*, *graphical* and *speech*. Thus, another features are activated producing a different UI showed at the right of the figure.

3 Validation of the Proposal

In order to validate our proposal, we have defined a scenario that illustrates how interaction can be adapted to provide an adequate obtrusiveness level in an Ambient Intelligence (AmI) environment. The adaptation takes into account different factors such as the location of mobile phone (at user's hand or nearby to the user), user location and message urgency. All these factors have an effect in the obtrusiveness level to be provided. Then, this scenario has been evaluated.

3.1 Smart Home Case Study

We applied our approach to a case study of a smart home environment based on the scenario of service adaptation we developed in [4]. We extended the services defined in the original case study in order to adapt the obtrusiveness level at which they are presented to the user.

The scenario of our case study describes a normal day in Bob's life and the way interaction mechanisms of different home services change depending on the context. Bob lives in a smart home with garden and a swimming pool. Every day, he gets up at 7 a.m. and drinks milk for breakfast while he watches a TV program before going to work. One day during breakfast, Bob runs out of milk. In reaction to this, the refrigerator added this item to the shopping list in an invisible manner for Bob. While he was watching the TV program, the system reminded him that he had an important meeting at work and he had to leave the house sooner. Therefore, the video service started to record it.

During the meeting, the smart home reminded Bob about watering the plants. Because of he had the mobile at hand, the notification appeared in a subtle manner suggesting him if he wanted that the system water the plants.

When he was going back to home, he was nearby of a supermarket and the mobile notified him about it, showing the map to arrive to the supermarket. When he was there, the map was changed by the floor map of the supermarket. At the same time, the mobile suggested him the items to the shopping list that were available in that supermarket. While Bob was buying, the mobile suggested him a television series to record. When he arrived at home, he put the mobile to charge. While it was charging, pool was cleaned automatically.

In this scenario, several services are presented at different obtrusiveness level. For example, video recorder service is presented first in a *reactive-invisible* manner because it begins to record automatically in reaction to the user leave. But then, the same service *proactively* notifies the user about to record the program in a *subtle* manner. In this way, users could evaluate the adaptation.

We developed a prototype version for the system described and conducted an experiment[1]. The experimental setup included an HTC Magic mobile device running Android Operating System. The experiment showed that by following our technique, services with the properly interaction mechanisms in terms of obtrusiveness can be obtained.

3.2 Evaluation Results

In order to evaluate the user acceptance of the system and determine whether the interaction has been adapted properly, we ran a controlled experiment over 15 participants. We used an adapted IBM Post-Study questionnaire [12] in conjunction with the questionnaire defined by Vastenburg et al. in [16] to evaluate home notification systems for three dimensions. This three dimensions were: *Usability of the system*, *messages acceptability* and *interaction adaptation*. The first dimension focuses on measuring users' acceptance with the usability of the system; the second one focuses on the general acceptability considering the messages and the user activity at the time of notification; and finally, the third dimension is about users' satisfaction in the interaction adaptation.

The study was conducted in our lab in order to simulate the different scenarios in which the experiment was based on. Users adopted Bob's role and perform the activities earlier described, filling the questionnaire at the end.

A total of 15 subjects participated in the experiment (6 female and 9 male). Most of them had a strong background in computer science. Participants were between 23 and 40 years old. 8 out of 15 were familiar with the use of a smartphone, and three own an Android device similar to the one used in the experiment. We applied a Likert scale (from 1 to 5 points) to evaluate the items defined in the questionnaire. Figure 4 shows a summarized table of the obtained results[2].

[1] Screenshots of the developed prototype are showed in http://www.pros.upv.es/labs/projects/interactionadaptation

[2] The complete dataset of the experimental results can be downloaded from http://www.pros.upv.es/labs/projects/interactionadaptation

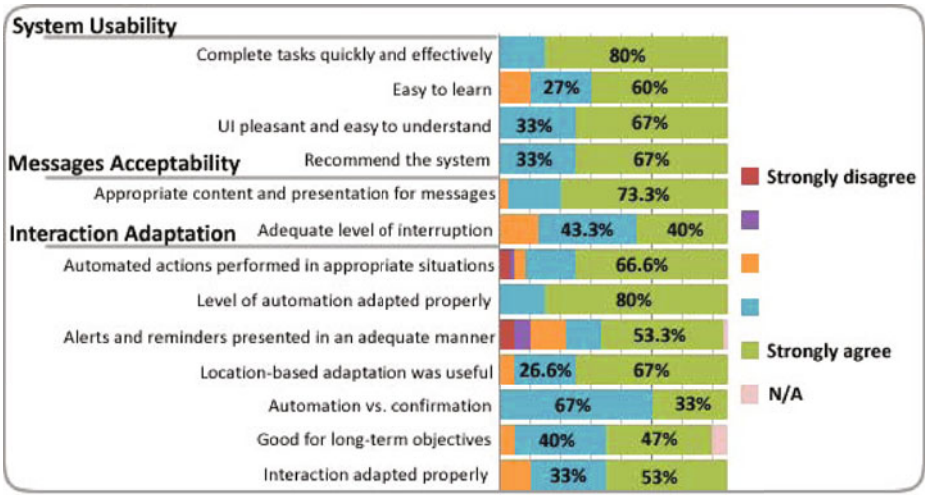


Fig. 4. Summarized results

More than 70% of the people strongly agreed that using the system they were able to complete the tasks and scenarios effectively and quickly. All users considered (4 or 5 points) the user interface to be pleasant and easy to understand.

With regard to the *messages acceptability*, the results were also positive, but more dispersion was found in them. This was due the different perception each user had about what was considered to be a relevant or urgent message. In the study made by Vastenburg et al. [16], they pointed out that the more urgent the message was considered to be, the higher the level of intrusiveness should be. In our results, the content and presentation of the different messages was considered appropiate by the 73% of the subjects. Some users (20%) found some services to be intrusive, but the interruption level was in general (80%) considered adequate.

Regarding the *interaction adaptation*, automated tasks outcomes are not always discovered (33% of subjects), but 80% of subjects strongly agreed in that automated actions had performed in appropiate situations. There were some exceptions that were suggested in the comments such as "I would like to receive the pool notification and can postpone it" or "When the system clean the pool do not inform the user about that". Although the adaptation provided was considered adequate (more than 80% considered it appropriate for all the services), most of the complaints were related to the level of control provided. Some users would like to be able to undo actions they are notified about such as the video recording, many (67%) did not considered watering the plants deserving a notification, and the suddenly change of the outdoor to an indoor map of the supermarket made some users (33%) feel they were loosing control.

The initial results obtained show that by following our approach we can adjust the obtrusiveness level for the services in a detailed manner. Nevertheless, additional experimentation would be required to analyse the adaptation during longer periods. Due to time constraints, we gave the users a script to follow to reproduce specific tasks and contexts of use. Using a script that was conformant to the process rules did not allow to evaluate the system in a more realistic context where services are competing with daily activities.

4 Related Work

This work deals with the modeling of interaction, taking into account its nomadic and context-aware nature. Calvary et al. in [3] give an overview of different modeling approaches to deal with user interfaces supporting multiple targets in the field of context-aware computing. Our proposal introduces the notion of features to relate them in a decoupled fashion. In this way, we can describe UI adaptation over multiple platforms (1) in a declarative manner and (2) taking into account different aspects such as obtrusiveness.

Different modeling languages are focused on the description of interaction. Aproaches such as UIML [1] or UsiXML [13] define domain-specific languages that are specifically designed from the beginning to deal with the description of user interfaces in a device-independent manner. All of these approaches consider pre-defined set of context factors and do not decompose them to exploit their commonalities and differences as our proposal does by means of the Feature Model. For example, a noisy context and a user with an auditory impairment require interaction not to be provided by means of audio. By considering the specification in terms of features (as opposed to specifying it for each context), both cases are expressed as the exclusion of the auditory feature.

Several works have studied home notifications systems, which is the area in which our approach was applied. These works focus on minimizing unnecessary interruptions for the user [14]. Many of the previous studies have concentrated on predicting the cost of interrupting the user by means of Bayesian reasoning models [10] or comparing different mental or task stages during which interruption occurs [9]. Given this background, our notifications took into account the context of the user in deciding the obtrusiveness level of each event.

5 Conclusions

The challenge in an environment full of embedded services is not only to make information available to people at any time, at any place, and in any form, but to reduce information overload by making information relevant to the task-at-hand [6]. Notification methods should achieve the right balance between the costs of intrusive interruptions and the loss of context-sensitivity of deferred alerts [10].

This work provides an approach to define services according to the context of each user in terms of obtrusiveness without duplicating efforts in the development. The mechanisms required to set and dynamically update the obtrusiveness level from contextual information fall out of the scope of the present work. Further work will be dedicated to develop tools to systematize the proposal enabling end-users to set their preferences and integrate it with the Model-based Reconfiguration Engine (MoRE) [4] to achieve a dynamic reconfiguration in response to context variations.

References

1. Abrams, M., Phanouriou, C., Batongbacal, A.L., Williams, S.M., Shuster, J.E.: Uiml: an appliance-independent xml user interface language. In: WWW 1999, pp. 1695–1708. Elsevier North-Holland, Inc., New York (1999)
2. Van den Bergh, J., Coninx, K.: Using uml 2.0 and profiles for modelling context-sensitive user interfaces. In: Proceedings of the MDDAUI 2005, CEUR Workshop (2005)
3. Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L., Vanderdonckt, J.: A unifying reference framework for multi-target user interfaces. Interacting with Computers 15(3), 289–308 (2003)
4. Cetina, C., Giner, P., Fons, J., Pelechano, V.: Autonomic computing through reuse of variability models at runtime: The case of smart homes. Computer 42(10), 37–43 (2009)
5. Czarnecki, K., Helsen, S., Eisenecker, U.: Staged configuration using feature models. In: Nord, R.L. (ed.) SPLC 2004. LNCS, vol. 3154, pp. 266–283. Springer, Heidelberg (2004)
6. Fischer, G.: User modeling in human-computer interaction. User Modeling and User-Adapted Interaction 11(1-2), 65–86 (2001)
7. Gibbs, W.W.: Considerate computing. Scientific American 292(1), 54–61 (2005)
8. Giner, P., Cetina, C., Fons, J., Pelechano, V.: Developing mobile workflow support in the internet of things. IEEE Pervasive Computing 9(2), 18–26 (2010)
9. Ho, J., Intille, S.S.: Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In: CHI 2005, pp. 909–918. ACM, New York (2005)
10. Horvitz, E., Koch, P., Apacible, J.: Busybody: creating and fielding personalized models of the cost of interruption. In: CSCW 2004, pp. 507–510. ACM, New York (2004)
11. Ju, W., Leifer, L.: The design of implicit interactions: Making interactive systems less obnoxious. Design Issues 24(3), 72–84 (2008)
12. Lewis, J.R.: Ibm computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int. J. Hum.-Comput. Interact. 7(1), 57–78 (1995)
13. Limbourg, Q., Vanderdonckt, J., Michotte, B., Bouillon, L., Lopez-Jaquero, V.: UsiXML: A Language Supporting Multi-path Development of User Interfaces (2005)
14. Ramchurn, S.D., Deitch, B., Thompson, M.K., Roure, D.C.D., Jennings, N.R., Luck, M.: Minimising intrusiveness in pervasive computing environments using multi-agent negotiation. In: First International Conference on Mobile and Ubiquitous Systems, pp. 364–372 (2004)
15. Tedre, M.: What should be automated? Interactions 15(5), 47–49 (2008)
16. Vastenburg, M.H., Keyson, D.V., de Ridder, H.: Considerate home notification systems: A user study of acceptability of notifications in a living-room laboratory. Int. J. Hum.-Comput. Stud. 67(9), 814–826 (2009)

A Dynamic Time Warping Approach to Real-Time Activity Recognition for Food Preparation

Cuong Pham, Thomas Plötz, and Patrick Olivier

Culture Lab, School of Computing Science, Newcastle University, United Kingdom
{cuong.pham,t.ploetz,p.l.olivier}@ncl.ac.uk

Abstract. We present a dynamic time warping based activity recognition system for the analysis of low-level food preparation activities. Accelerometers embedded into kitchen utensils provide continuous sensor data streams while people are using them for cooking. The recognition framework analyzes frames of contiguous sensor readings in real-time with low latency. It thereby adapts to the idiosyncrasies of utensil use by automatically maintaining a template database. We demonstrate the effectiveness of the classification approach by a number of real-world practical experiments on a publically available dataset. The adaptive system shows superior performance compared to a static recognizer. Furthermore, we demonstrate the generalization capabilities of the system by gradually reducing the amount of training samples. The system achieves excellent classification results even if only a small number of training samples is available, which is especially relevant for real-world scenarios.

1 Introduction

Ambient Assisted Living (AAL) is a key domain of Ambient Intelligence (AmI), which focuses on the development of technology that supports humans during their activities of daily living (ADL). Applying such technology to people's private homes offers, for example, a realistic opportunity for age-related impaired people to live more independent and longer in their homes. Especially the kitchen plays an important role in people's life, as it is an indispensable place where many everyday life activities such as cooking and food preparation take place. Although the AmI research community has recently made a stride in real-world applications for health care at private homes, the technology driven support for cooking and food preparation activities is still relatively under-explored.

Technology support in the kitchen would include nutritional advice for, e.g., recipe selection and automatic guidance whilst cooking. Especially the latter is of major importance for people with age-related impairments, such as stroke and dementia. For these people it is often difficult to focus on the particular kitchen task and, hence, to complete the cooking. Monitoring kitchen activities and eventually providing situated advice would help substantially, e.g., for following a recipe and would thus give these people more independence in their private homes.

As a first step towards automatic kitchen assistants we developed an activity recognition system for kitchen environments that automatically recognizes relevant food preparation tasks [9]. Accelerometers were embedded into kitchen utensils and sensor

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 21–30, 2010.

data are recorded whilst the utensils are being used. Recognition was based on a decision tree classifier that processes fixed length statistical feature vectors extracted in a sliding window procedure. By means of a closed set evaluation we demonstrated the reliable recognition of ten common kitchen tasks, which represents the proof of concept for the overall approach.

In this paper we substantially improve our AR system for the kitchen by integrating dynamic time warping (DTW) based classifiers. The motivation is to adapt Activity Recognition towards the users' idiosyncrasies by means of an automatic maintenance of a proper template database, which is used for the actual recognition. DTW based classification aligns sequential data in a way that preserves potentially existing internal structures of the data, which is beneficial for the analysis of real-world time-series. Compared to [9], in this paper we apply the system to a more realistic evaluation tasks with an open inventory of kitchen activities to be recognized, i.e., integrating segmentation of continuous data streams. Furthermore, we investigate the effect smaller training sets have on the accuracy of the recognition system. Especially the latter is relevant for realistic tasks where the amount of labeled sample data is usually limited since manual annotation is tedious and labor intensive, hence costly. The whole system is implemented as a real-time recognition system, which is integrated into [9], our lab-based pervasive kitchen environment.

2 Related Work

Technology-augmented kitchens have been used to explore the application of AmI in the home. For example, the AwareKitchen was an intelligent environment equipped a number of sensors such as microphone, forces, accelerometer etc. to detect cutting food activity[1]. CounterIntelligence[2], an augmented reality kitchen, is able to support the users while they are performing cooking task by displaying instructive text. An application of RFID technology embedded in a kitchen counter was described in [6]. A number of other design proposals relate to the provision of situated advice on food and cooking (cf., e.g., [7]). Moreover, with the goal of helping old people to live more independent in the homes, the Ambient kitchen, a design of situated services using a high fidelity prototyping environment in which technologies such as wireless accelerometers, RFID, IP cameras, projectors are completely embedded into the environment, is proposed in [8].

A significant amount of previous work on activity recognition is based on sensors wore on user's body and the application of pattern recognition techniques. In [11], for example, a Dynamic Bayesian Network approach was described that aims at recognizing high-level household activities based on object use.[12] explored hidden Markov models (HMMs) for activity modeling and recognition. In [9], an AR framework for recognizing low-level typical food preparation activities using accelerometers and classical machine learning approaches has been presented.

Dynamic Time Warping (DTW) is a standard technique for the comparison of time series data [15]. The key idea is to align two sequences in an optimal way, i.e., minimizing "costs" for this "warping" (see Sec. 3.1). Numerous applications of DTW have been developed being as diverse as personalized gesture recognition [10], speech recognition [15], and hand printed signature verification [16].

3 Dynamic Time Warping Based Activity Recognition for Food Preparation

Aiming to support people in their activities of daily living, the focus of our research is on the development of a kitchen monitoring system that ultimately will provide automatic situated assistance for kitchen tasks like food preparation. The basis for such a system is reliable and efficient activity recognition. For this purpose we proposed to integrate accelerometers into kitchen utensils, which are used for food preparation [9]. Analyzing the data recorded by the sensor-equipped utensils allows to recognize typical food preparation activities in real-time, i.e., while the person is acting in the kitchen. Based on this, situated support for food preparation becomes possible.

Analyzing food preparation activities in more detail, it becomes clear that even the most fundamental activities exhibit substantial variance depending on the personal preferences of how to handle the food ingredients and the utensils. Consider, for example, the process of shaving a carrot using a knife. Some persons perform long, slow movements of the utensil along the carrot towards themselves, which is comparable to "carving" the vegetable. Others tend to perform short, fast cuts of the carrot's surface thereby using the knife more in a "chopping" way. Although both kinds of movements differ substantially, they represent the same kind of activity and an automatic recognition system needs to cope with it.

In order to deal with the aforementioned users' idiosyncrasies, we developed a fully automatic, real-time activity recognition system, which is outlined in Fig. 1. Accelerometer data are recorded while a person is working in the kitchen. As in our previous work, sensor equipped standard kitchen utensils integrating modified Wii remotes are used. For continuous sensor data streams (x,y,z) with a sampling frequency of 40Hz, then frames of 64 contiguous samples are extracted in a sliding window procedure (50 percent overlap; lower part of Fig. 1). Following some basic preprocessing and trivial movement detection (simple threshold based procedure), then the actual classification of the extracted frame regarding the activities of interest is performed (middle part of the figure). This recognition procedure is performed as a DTW-based template comparison with an automatically maintained template database. This database contains representative templates for the activities of interest together with activity specific thresholds for acceptance / rejection (upper right part). By means of the analysis of the DTW scores the template database is continuously adapted to represent the idiosyncrasies of the particular activities performed by different users. The output of the system consists of classification hypotheses for every extracted frame including possible rejection, which effectively means segmentation of continuous sensor data streams.

In the following we will first briefly summarize the theoretical foundations of Dynamic Time Warping before giving detailed descriptions of the key components of the overall recognition system.

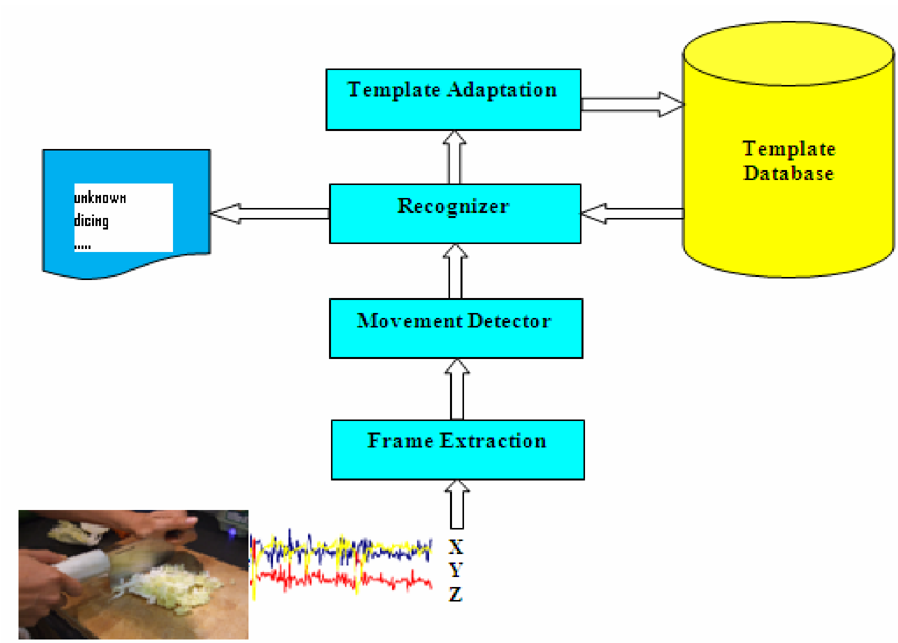


Fig. 1. Activity Recognition for Food Preparation Tasks – System Overview

3.1 Dynamic Time Warping – A Brief Overview

DTW (cf. [15]) is a constrained, non-linear pattern matching method based on dynamic programming to measure the dissimilarity between two time series. Let $O = o_1,o_2,..,o_m$ and $Y = y_1,y_2,..,y_n$ be two time series. DTW finds an optimal mapping from O to Y by reconstructing a *warp* path W that optimizes the maping of the two sequences: $W = w_1,w_2,\ldots,w_K$ where $\max\{m,n\} \leq K < m+n$, with K denoting the length of the warp path. The warp path constrained in the sense that it is anchored by the start and end points of both sequences. To map the in-between elements of both time series a step-wise distance minimization for every position is performed:

$$\delta(W) = \min \{ \sum_{k=1}^{K} \delta(w_{ki}, w_{kj}) \} \tag{1}$$

The actual distance calculation is usually (but not necessarily) based on the Euclidean distance $d(o_i,y_j)$ and dynamic programming [15]:

$$\delta(i,j) = d(o_i,y_j) + \min \{ \delta(i-1,j), \delta(i-1,j-1), \delta(i,j-1) \} \tag{2}$$

where i and j represent monotonically increasing indices of the time series O and Y. The resulting matching cost $\delta(W)$ is then usually normalized to the range of [0,1] to ensure comparability.

3.2 DTW-Based Recognition

By means of the accelerometers integrated into the kitchen utensils continuous three-dimensional (x,y,z) data streams are recorded. In a sliding window procedure 64 contiguous samples are summarized to frames. Adjacent frames overlap by 50 percent. The actual parameterization of the frame extraction process has been optimized in previous experiments (cf. [9]).

For every observation frame O[u], which is recorded for a particular utensil u, activity recognition is performed using the DTW-based algorithm as outlined in Fig. 2. Note that by means of a trivial threshold comparison only those frames are considered where the utensil was actually moving. After computing and sorting the DTW scores for all templates (lines 4:12 in Fig. 2), the set of the smallest min(K,n) scores is compared to the activity-specific thresholds (lines 13:19). If none of the DTW score s contained in the sorted list `cost` is smaller than the particular threshold the observation frame O is rejected, i.e., assigned to the unknown class. Heuristically we chose K=10, which provides reasonable results for acceptance / rejection on a cross validation set. The Threshold function (`Thresh,` line 14) retrieves the class-based threshold of the template Y[index[i]].

```
Input:      An observation frame O, utensil u, number K of
            sorted match scores to analyze for final result
Output:     Activity hypothesis
            //Extract templates
1:          CurrentTemplates= ExtractTemplateDB(u);
2:          n=the number of templates in CurrentTemplates;
3:          For i from 1 to n do
4:              Y[i]=CurrentTemplate.template[i];
5:              cost[i]= DTW(O, Y[i]);
6:              index[i] = Y[i].Id;
7:          End for
            // sort and maintain indices of the templates
8:          Sort(cost, index);
9:          For i from 1 to min{K,n} do
                //Acceptance
10:             If cost[i]<Thresh(Y[index[i]]) then
11:                 activity(Y[index[i]])→activity_list;
12:                 break;
13:             Else
                //Rejection
14:                 unknown activity → activity_list;
15:             End if
16:         End for
17:         Return (activity_list);
```

Fig. 2. DTW-based activity recognition

The recognition procedure utilizes activity specific thresholds T_a. Applying Chow's rule [13] to our domain, an observation frame o is classified as being a particular activity a if:

$$\delta(o,t) = \min_{j=1..N} \{ \delta(o,t_j)\} < T_a \tag{3}$$

where t represents a template from the database, which represents the activity a. N is the overall number of templates in the database. A frame is classified as unknown, i.e., being rejected, if:

$$\delta(o,t) = \min_{j=1..N} \{ \delta(o,t_j)\} \geq \max_{i=1..n} \{T_i\} \tag{4}$$

where n denotes the number of classes of interest. The class-based thresholds were manually selected through a 5-fold cross validation procedure.

3.3 Template Adaptation

In order to adapt the overall recognition system towards users' idiosyncrasies, the template database is continuously being updated, i.e., templates are removed and added if necessary. The adaptation scheme used can be described as follows. Let f_k define the weighted histogram of recognition hypotheses for a particular time step k consisting of activity specific entries $f_k(a)$, where specifies the particular activity, and w denotes a (heuristically chosen) adaptation weight:

$$f_k(a) = \begin{cases} f_{k-1}(a) \cdot w & \text{if} \quad \delta(a,t) < \alpha \\ f_{k-1}(a) \cdot (1-w) & \text{otherwise} \end{cases} \tag{5}$$

α denotes an acceptance / rejection threshold, which is derived from the activity specific template thresholds: $\alpha = T_a/(1+T_a)$. All thresholds were optimized in a separate cross-validation procedure. Let $\gamma_k(a)$ be the cumulative number of recognitions of activity a at time k. The positive probability of activity a at time k $\rho_k(a)$ is computed as:

$$\rho_k(a) = \frac{f_k(a)}{\gamma_k(a)} / \sum_k \frac{f_k(a)}{\gamma_k(a)} \tag{6}$$

The negative probability of activity a at time k is hence defined as:

$$\varphi_k(a) = \frac{1-\rho_k(a)}{\sum_k 1-\rho_k(a)} \tag{7}$$

At time k, if an observation frame makes $\varphi_k(a) = \min\{\varphi_i(a)\}$ for i=1..k, then this frame will be updated as a template into the *positive* list in the template database for later use (adaptation). The template which makes $\varphi_i(a) = \max\{\varphi_i(a)\}$for i=1..k is moved to the *negative* list in the template database at the same time. The negative list is used only when recognizer has recognized an unknown activity on the positive list, then recognizer does one more try on the negative list before returning the activity list.

4 Experimental Evaluation

In order to evaluate the applicability of the proposed DTW-based approach to activity recognition, we performed a number of practical experiments. Therefore, we used the (publically available) dataset from our previous work [9], which covers 20 persons pursuing typical food preparation tasks (salad and sandwich making) using our sensor-equipped utensils. No further constraints were given. Ten typical low-level activities were subject to recognition, namely *chopping*, *peeling*, *slicing*, *dicing*, *scraping, shaving*, *scooping*, *stirring*, *coring, spreading.* Additionally a considerable amount of sensor data, which does neither belong to one of the ten known activities nor to "idle" is included in the dataset. In total more than 6 hours of sensor data have been collected from four sensor equipped kitchen utensils (knifes and spoon).

Extending the proof-of-concept, which was given in [9], we aimed for realistic experiments using the recognition system in real-world scenarios. This implies that the recognition system was applied online, i.e., continuous data streams had to be segmented and classified (open lexicon with rejection) in real-time, i.e., with negligible latency (results given in Sec. 4.1). Furthermore, we are interested in the dependency of the recognition procedure on the number of annotated samples available for training. Since manual annotation is tedious and costly, it is somewhat unrealistic to rely on large training sets for setting up the recognition system. Consequently, we performed a second set of experiments where the number of training samples has been decreased step by step. Classification results are reported in section 4.2.

Recognition results are reported as frame-wise precision and recall values. The precision for some activity a was calculated by dividing the number of correctly classified frames by the total number of frames classified as being a (i.e. true positives/(true positives + false positives)). Recall was calculated accordingly as the ratio of the number of correctly classified frames of a and the total number of frames of a (true positives/total number of frames of A). Baseline results for the evaluation are given by the decision tree based system described in [9].

4.1 Results for Full Training Set

For the first set of experiments we used all available training data (see below) to estimate the recognition system. When the dataset was recorded, it was manually annotated by 3 independent subjects. The consensus of the three annotators serves as ground truth. Since we recorded complete cooking sessions the dataset for obvious reasons is dominated by idle "activities", i.e., frames recorded while the particular utensil of interest is not moved at all. A quick subject-independent test including all idle frames led to 96.36% overall accuracy as the recognition of "idle" is almost perfect (see Sec. 3: simple threshold comparison). To avoid this over-optimistic and not quite informative evaluation, we limited the set of "idle" frames to four per utensil, which were randomly selected per subject. This effectively truncates the dataset to 12,265 frames (of 64 samples each).

The evaluation was performed in a "leave-one-subject-out" manner, i.e., we trained the recognizer using the data from 20 subjects, and tested on those data recorded for the remaining subject. This process was repeated for all 21 subjects and results were averaged. The overall performance of the proposed DTW-based approach are presented in

Tab. 1. Additionally the results for the baseline system (Decision Tree C4.5) are given. It can be seen that the new approach clearly outperforms the baseline with overall precision rates of approx. 83% (CBT-DTW) vs. 77.9% (DT C4.5), and recall rates of 82.8% (CBT-DTW) vs. 76.7% (DT C4.5). All differences are statistically significant. Additionally Tab. 2 illustrates the aggregated confusion matrix for the subject-independent evaluation of CBT-DTW.

Table 1. One-subject-leave-out evaluation (all figures in percent)

Activity	CBT-DTW			Decision Tree C4.5		
	Precision	Recall	False Positive	Precision	Recall	False Positive
chopping	82.61	88.54	2.22	82.21	87.5	7.37
coring	77.02	81.94	0.21	74.02	77.7	4.12
dicing	51.16	54.63	0.25	24.87	18.7	4.25
peeling	72.76	80.63	0.53	88.7	95.9	3.91
scraping	80.09	81.1	0.12	56.8	56.3	3.37
shaving	72.79	82.73	0.28	55.11	59.7	2.91
slicing	70.31	70.73	1.21	33.47	26.6	4.95
spreading	71.06	86.57	0.77	54.33	44.4	2.32
scooping	97.92	94.55	0.78	91.2	86.3	2.6
stirring	84.77	86.98	0.08	81.63	85.92	1.26
idle	100	100	0	100	100	0
unknown	91	80.92	4.96	85.3	83.2	9.82
Overall	**83.02 ± 4.8**	**82.78 ± 5.5**	**2.61 ± 1.03**	**77.9 ± 8.7**	**76.7 ± 6.5**	**6.29 ± 2.1**

Table 2. Aggregated confusion matrix for one-subject-leave-out evaluation for CBT-DTW [%]

	a: chopp.	**b: coring**	**c: dicing**	**d: peel.**	**e: scrap.**	**f: shav.**	**g: slic.**	**h: spread.**	**i: scoop.**	**j: stir.**	**k: idle**	**l: unkn.**
A	88.4	0.14	5.5	0	0.3	0	4.56	0	0	0	0	1.09
B	0	81.9	0	6.4	3.1	0	0	38.9	0	0	0	4.72
C	36.6	0	54.6	0	1.2	0	5.2	0	0	0	0	2.48
D	0	4.0	0	80.6	5.8	3.7	0	0	0	0	0	5.9
E	2.3	1.06	0.1	4.05	81.1	2.6	0.96	0	0	0	0	7.81
F	2.0	0	0	8.43	1.21	82.7	0	0	0	0	0	5.62
G	15.0	1.71	7.63	0	1.37	0	70.7	0	0	0	0	3.53
H	0	0	0	0	5.76	1.68	0	86.6	0	0	0	5.99
I	0	0	0	0	0	0	0	0	94.6	2.08	0	3.38
J	0	0	0	0	0	0	0	0	8.07	87.0	0	4.95
K	0	0	0	0	0	0	0	0	0	0	100	0
l	3.28	0.92	0.03	2.3	2.96	0.58	1.4	3.51	3.18	0.95	0	80.9

4.2 Results for Reduced Training Sets

In the second set of experiments we analyzed the dependency of the classification accuracy on the number of samples available for training the recognizer. We therefore gradually reduced the amount of annotated training samples by randomly selecting frames to be excluded from training and evaluated the resulting recognizers on the remaining data.

Fig. 3 illustrates the classification accuracies of both the proposed CBT-DTW approach and the baseline system (DT C4.5). The number of frames used for model training is given at the x-axis, whereas the y-axis represents the classification accuracies. For the sake of clarity in the illustration the (discrete) figures are connected to continuous curves. It can be seen that the proposed DTW-based recognition approach greatly generalizes even if only small amounts of training data were available. For example, with only 5 labeled frames for training (per activity) the accuracy of CBT-DTW is still about 79%. The performance of the baseline system here drops to approximately 55%. Fig. 4 also illustrates the general superiority of the proposed approach compared to the baseline system.

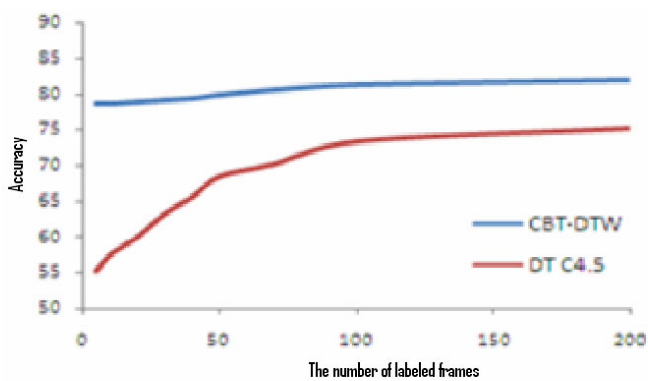


Fig. 3. Classification accuracies in dependence of the amount of training data

5 Summary

Automatically monitoring food preparation activities is a key for the development of kitchen assistance systems that, for example, provide situated cooking support for people with age-related impairments like dementia. For non-intrusive activity recognition we integrated accelerometers into kitchen utensils and analyze the sensor data recorded while people cook using these enhanced utensils.

In this paper we presented an activity recognition system that automatically adapts towards the idiosyncrasies of people using kitchen utensils. Based on a Dynamic Time Warping procedure a template matching system has been developed, which successfully segments and recognizes ten low-level kitchen activities by analyzing contiguous frames of sensor readings. The adaptation of the AR system towards variants of certain activities is pursued by an automatic maintenance procedure, which effectively updates the template database if necessary.

By means of an experimental evaluation on a large, realistic datasets that covers unconstrained food preparation we demonstrated the capabilities of the proposed approach. In a second set of experiments we gradually reduced the number of training samples. The proposed DTW-based recognition system shows superior generalization even if only a few samples are available for training.

References

1. Kranz, M., Schmidt, A., Rusu, B., Maldonado, A., Beetz, M., Hornler, B., Rigoll, G.: Sensing Technologies and the Player-Middleware for Context-Awareness in Kitchen Environments. In: Proc. 4th Int. Conf. on Networked Sensing Systems (2007)
2. Bonanni, L., Lee, C.H., Selker, T.: CounterIntelligence: Augmented Reality Kitchen. In: Proc. CHI 2005, pp. 2239–2245 (2005)
3. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In: Proc. Int. Conf. Pervasive Comp., pp. 1–17 (2004)
4. Ravi, N., Dandekar, N., Mysore, P., Littman, M.L.: Activity recognition from accelerometer data. In: Proc. Conf. Innovative Applications of Artificial Intelligence (IAAI), pp. 1541–1546 (2005)
5. Chang, K.-H., Liu, S.-Y., Chu, H.-H., Hsu, J., Chen, C., Lin, T.-Y., Huang, P.: Dietary–aware dining table: Observing dietary behaviors over tabletop surface. In: Proc. Int. Conf. Pervasive Comp., pp. 366–382 (2006)
6. Chi, P.-Y., Chen, J.-H., Chu, H.-H., Chen, B.-Y.: Enabling nutrition-aware cooking in a smart kitchen. In: Proc. CHI 2007 Extended Abstracts on Human Factors in Computing Systems, pp. 2333–2338 (2007)
7. Tran, Q.T., Calcaterra, G., Mynatt, E.D.: Cooks collage: Deja vu display for a home kitchen. In: Proc. HOIT 2005 Conf. on Home-Oriented Informatics and Telematics, pp. 15–32 (2005)
8. Olivier, P., Monk, A., Xu, G., Hoey, J.: Ambient Kitchen: Designing situated services using a high fidelity prototyping environment. In: Proc. Int. Conf. Pervasive Technologies Related to Assistive Environments (2009)
9. Pham, C., Olivier, P.: Slice&Dice: Recognizing Food Preparation Activities using Embedded Accelerometers. In: Proc. Europ. Conf. Ambient Intell., pp. 34–43 (2009)
10. Liu, J., Zhong, L., Wickramasuriya, J., Vasudevan, V.: uWave: Accelerometer-based personalized gesture recognition and its applications. In: Proc. Int. Conf. Pervasive Comp. and Comm., pp. 1–9 (2009)
11. Wu, J., Osuntogun, A., Choudhury, T., Philipose, M., Rehg, J.: A scalable approach to Activity Recognition based on Object Use. In: Proc. Int. Conf. Comp. Vision (2009)
12. Wang, L., Gu, T., Tao, X., Lu, J.: Sensor-Based Human Activity Recognition in a Multiuser Scenario. In: Proc. Europ. Conf. Ambient Intell., pp. 78–87 (2009)
13. Chow, C.K.: An Optimal Error and Reject Tradeoff. IEEE Trans on Information Theory 16(1), 41–46 (1970)
14. Xi, X., Keogh, E., Shelton, C., Wei, L.: Fast Time Series Classification Using Numerosity Reduction. In: Proc. Int. Conf. Machine Learning (2006)
15. Myers, C.S., Rabiner, L.R.: A comparative study of several dynamic time-warping algorithms for connected word recognition. The Bell System Technical Journal 60(7), 1389–1409 (1981)
16. Jayadevan, R., Kolhe, S.R., Patil, P.M.: Dynamic Time Warping Based Static Hand Printed Signature Verification. Journal of Pattern Recognition Research 4(1), 52–65 (2009)

Refining Interaction Designs through Simplicity*

Pablo Muñoz, Pau Giner, and Vicente Pelechano

Centro de Investigación en Métodos de Producción de Software
Universidad Politécnica de Valencia
Camino de Vera s/n, 46022 Valencia, España
{pmunoz,pginer,pele}@pros.upv.es

Abstract. With more and more devices in our surroundings, users increasingly-consume applications and digital services which compete for their attention. Therefore, users appreciate simplicity when they interact with them because a less intrusive interaction allows users focus on the task at hand. This work relates an approach based on the concept of simplicity. We propose an iterative interaction design process in which we include a new role. This role is only in charge of providing simplified solutions from original designs. Moreover, we put on practice our proposal in a case study in which we design an application to support mobile workflows. We relate our experience and how new ideas were produced towards the definition of refactored interactions by means of the key of Simplicity.

Keywords: interaction, simplicity, mobile devices, user-centered design.

1 Introduction

Increasingly, we are surrounded by more and more devices. Devices such as laptops, smartphones, tablet PCs and set-top-boxes among others provide access to digital services but they also compete for user attention each day. For such an environment there is a challenge in providing users with their desired functionality but doing so in a way that intrude the users' mind as little as possible in order not to be regarded as obnoxious [2].

This work is focused on mobile devices since they are used in a changing context full of interruptions [3]. Considering the limitations (e.g., screen size, battery, etc.) and the advanced interaction capabilities of mobile devices (e.g., speech synthesis, multi-touch screens), the interaction design affects the degree in which users are able to complete their task effectively [4].

In this work we present an approach that is based on the concept of simplicity.Simplicity is about subtracting the obvious and adding the meaningful [6]. This principle is behind the design of successful products such as the iPod, a device that does less but costs more than other digital music players [6]. The use of simple interactions fits perfectly the mobile device characteristics since (1) optimizes their low

* This work has been developed with the support of MICINN under the project SESAMO TIN2007-62894 and co-financed with ERDF.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 31–40, 2010.

resources, and (2) hides the complexity of dealing with multiple interaction mechanisms and contextual information. These two factors contribute to focus on the task at hand.

Many usability guidelines try to aid users in fulfilling their goals effectively, efficiently and successfully in each context of use [1]. In order to obtain simpler designs, designers normally apply heuristic analysis to their interaction designs. There are different works describing heuristics or guidelines for general purpose design such as [8], [9]. There are also references related to touch gestures such as [10], for voice user interface design [11], or for mobile design [3]. In many cases simplicity is behind some of these guidelines. Nevertheless, by considering simplicity as an independent aspect for the system, some of the limits that reduce the design space are eliminated [7].

The contribution of this work is an iterative process for the interaction designbased on simplicity. In detail, we propose using simplicity as an orthogonal aspectby including a new role called "*simplifier*" who is focused on the task of simplifying the interaction. So, the commitments of this role are (1) to find complexity problems in interaction designs and (2) to provide solutions to face with these problems. This role is inspired by the Laws of Simplicity of Maeda [6]. These laws are metaphors applicable to different areas such as design, business, technology and life. In our approach, we interpret some of these metaphors to improve the interaction designs.

In addition, with the intention of evaluating our approach in practice we applied our interaction design process designing a contextual to-do list prototype for a mobile platform. In this work, we relate our experience applying our approach and describing which ideas were produced in the different iterations towards the definition of more user-centered interfaces by means of the key of the simplicity.

The remainder of this paper is structured as follows. Section 2 relates our approach for the interaction design process. Section 3 introduces our experience in applying our approach in the design of a workflow mobile application and section 4 draws conclusions.

2 The Interaction Design Process

In practice, we have found that designers find hard to simplify their own solutions. A designerthat produces a solution has a deep understanding of this solution which makes it hard to detect which are the aspects that need to be simplified. So, designers want to accomplish all the users' demands, but too much attention to the needs of the users can lead to a lack of cohesion and added complexity in the design [12]. Furthermore, the lack of standardization affects the practice in interaction design research [5]. This fact provokes that designers normally were disoriented when they create designs. Moreover, constraints such as the size of the screens or the successive changes in the user's context add more complications.

In order to face up to these problems we propose to introduce an iterative process for the interaction design based on simplicity. In detail, the process consists in refining the designs successively in different iterations by means of two involvedroles: the designer role which creates the interaction designs and thesimplifier rolewhich proposes advices to simplify them.

Only the designer is aware of the customer's demands and it creates designs that cover them. Nevertheless, in order to avoid bias in designs, the simplifier is not aware of the particular requirements for the project. It only is in charge of finding complexities in designs and proposing simplified solutions.

2.1 Giving Simplicity Advices by Means of the Laws of Simplicity

Going into details about how the simplifier gives its advices, the proposed solutions are provided by the useof the Laws of Simplicity of Maeda [6]. These laws are metaphors adaptable to different areas such as design, business, technology or life. Although they are not specific guidelines, they provide reflections and experiences based on keeping Simplicity.Thereby, weinterpreted the laws that we considered useful to fix problems in interaction design and we defined seven specific guidelines-forsupportingthe simplifier in giving advices based on simplicity.Specially, the first four guidelines have the goal of determining and organizing the information that must be showed with the purpose of avoiding overwhelmed designs. However, the rest have the goal of providing enhancements such as intuitiveness or good feelings over these selected information.

In detail, we inspired by the following laws: *"reduce"*,*"organize"*, *time"*, *"learn"*, *"context"*, *"emotion"*and*"the one"*.So, by means of these laws we defined the next guidelines:

1. **Remove redundant information or irrelevant data.** Check if a design contains removable information such as irrelevant data or redundant information. In this case remove it.
2. **Select the essential.** When there isn't redundant information, select the most important information and add it to the design.
3. **Complement and contextualize the essential.** After selecting the essential, complement it by taking into account the current contextual information and adapt the showed information according to this.
4. **Organize the secondary.** After complementing the essential, group and organize the rest of the functionality which isn't removable. Then, relocate it in a secondary level for don't affect the essential.
5. **Add more intuitiveness.** Move popular knowledge to the application in order to provide a more intuitiveness interaction.
6. **Save time.** Save time as much as possible without affecting the essential with the purpose of users complete its activities more fluently.
7. **Generate emotions.** Achieve users have good feelings which lead to a satisfying interaction.

Regarding how use these guidelines, they can be applied one by one or collectively. Furthermore, the use of all guidelines is not required since a good design can't need some guidelines.

2.2 Following the Process Step by Step

In order to refine the designs successively, the designs are refactored through different iterations. Going into details, a particular iteration contains the next four stages:

1. First, designer performs the designs according to the requirements.
2. Second, the simplifier applies the simplicity guidelines detectingcomplexities which it considered that could be simplified.
3. Third, the simplifier discusses with the designer until they reach an agreement about how redesign the designs.
4. Fourth, the designer refines the designs and meets the simplifier again.

Thus, the process iterates until the simplifier finally validates the designs. Thereby, the continuous feedback through different iterations leads to new designs which are refactored into simpler solutions. Figure 1 represents the proposed iterative process.

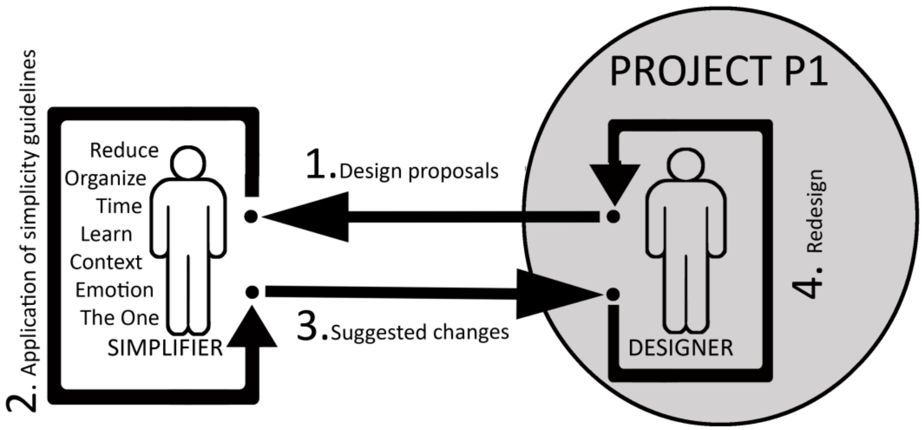


Fig. 1. Interaction design process

However, in our experience we found that simplification is not a straightforward process and requires taking many design decisions. Furthermore, it must be proved that the simplifier's advices don't remove value from the product and really reach simplicity. So, if its proposals are performed from hypotheses about the user, and since the simplifier is not aware of the requirements, the designer is whoshould verify whether it is true or nor with the final user (e.g., by prototyping).

Next, the following section describes our experience applying our approach for the design of a workflow supporting system.

3 Applying the Interaction Design Process to a Particular Case

In order to put in practice our approach we have developed a case study where two members of our research group played the two different roles. Thus, by means of this separation we avoided bias in the simplification process. On the one hand, the first played the designer role. Note that, since there wasn't any customer who performs demands in our case study, the requirements were initially established by the designer. On the other hand, the second played the simplifier role and it only had the goal ofsimplifying the designer's designs by using the defined guidelines to reach simplicity.

Regarding to the application target, we designed a mobile workflow application which allows users to manage a to-do list. This application covers the user's needs when they organize and complete their tasks on the go. For this purpose, we took into account three important attributes for organizing tasks: *location, time and priority*. So, according to these attributes we provided three viewswhich represent the tasks through these attributes. In detail, these three views are:

- *Cloud View.* Itvisualizes the tasks according their priority.
- *Timeline View.* It represents the tasks through the progress of the time.
- *Map View*. It provides the tasks over a map in a geo-located way.

Beyond how information is represented in the visualizations, other requirements were to provide support for the following operations: *complete task*, *edit task*, *delete task* and *send task to a contact*. Thus, designer was in charge of solving all of these requirements in its designs.

3.1 The Original Designs

The process of improving the interaction designs starts when the designer proposes its interaction designs according to the requirements. In our case, the initial designs provided by the designer which support these requirements were:

- **Cloud View.** This view consisted on showing the tasks as a cloud of tags. In order to differentiate the tasks according to the priority level, each tag was represented with a size and color font associated with its priority level.Regarding to the interaction with a concrete task, designer defined that when the user selects a task could be showed an overlay in the bottom of the screen. This overlay, showed in figure 3, included some data and operations.Specially, it showed the representative icon of the task, the name of the task, one button for each supported operation (*complete task, delete task, edit task, send task, view location*), priority labels that allow change priority and a button to confirm the changes.
- **Timeline View.**This view consisted on a horizontal and scrolled view guided bythe progress of the time. This view, showed in the third image of the figure 2, containsa collection of buttons for task operations at the top of the screen. Immediately below, tasks were represented by columns in which each column represented a task by its representative icon surrounded by a border that was referredto their priority. Each column also represented the established mean of transport and the associated costs of the task based on the required time and money. Furthermore, all icons were linked by directed arrows that expressed the order between tasks andat the bottom of the screen were viewed the hours when the tasks were programmed. Finally, about the interaction with a task, it was defined that user had to select a task and push the button according the action that the user want to perform.
- **Map View.** The initial design for the "*MapView*" represented each task in a geo-located way by its representative icon as we see in the second image of the figure 3. This icon was positioned inside a marker that represented the task priority by means of color use. Furthermore, the map also contained"*zoom in*" and "*zoom out*" controls to determinate the specified level of abstraction in the map. Regarding to

the interaction with a concrete task, the designer defined that when the end-user selects a marker an overlayis showed at the bottom of the screen. This overlay contained the same information that was included in the overlay related in "*Cloud View*".

Fig. 2. Initial designs: Cloud View, Map View and Timeline view

3.2 The Simplifier's Interventions

After the designer finishes the initial designs, the simplifier intervenes in the process. At this moment, the simplifier applies its simplicity guidelines and discusses with the designer how the designs could be simplified by means of the simplicity principles.

Briefly, in our experience the simplifier detected three general complexity problems: *the lack of context, too much components along the screen and the lack of intuitiveness.*Furthermore, the simplifier also detected some secondary problems such as the *waste of user's time* or *the irrelevant information showed*. These problems usually lead to users were overwhelmed and don't find easy to understand what they view, so that, they have problems to do its objectives.

Thereby, the simplifier made different proposals through different iterations in which it followed its guidelines with the purpose of solving these problems. Next, we summarize them according to the guideline that is used:

Removing redundant and irrelevant data. The goal in this point is to avoid overwhelming users by avoiding irrelevant data. In our particular study case, thesimplifier found in the initial designs some redundant information in the "*Timeline View*". In detail, it found irrelevant data such as the cost of performing a task like clean the diner room, which is showed also when itsvalue is zero as we can see in the third image of the figure 2.

Moreover, the designer included the redundant operation such as “view location”, which according to the designer is reachable from the menu and from a button represented in all views.

Select the essential. The goal of this point is to give special attention to the essential in order to users perform its main operations fluently. In detail, the simplifier found that in all views the most important operation was the “*complete task”* operation, but in each view it is always surrounded by a collection of buttons composed of less important operations which overwhelms users.So, the simplifier detected the need of modify these parts in designs which huddled different components in a small set of pixels.With this purpose, the simplifier proposed to maintainonly the “*complete task*” operation in the three views.

Organize the secondary. The goal of this point is (1) group the secondary and (2) relocate it in a second level. In detail, the simplifier group the operations “*edit task, delete task* and *send task to a contact*”. Thus, it proposed to move them to a more detailed view of the task providing an operation to navigate from the main context to this detailed view. So, interaction with a task changes as figure 3 shows.

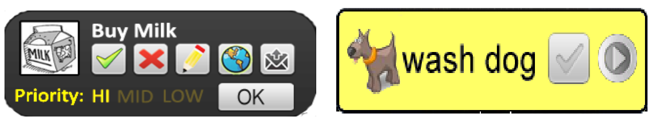


Fig. 3. Changes in the interaction with a task after the simplifier’s intervention

Complement and contextualize the essential. By complementing and contextualizing the essential we aim to provide users of a better understanding about the current context with the purpose of facilitating they complete their activities.

Specially, in our study case, the simplifier detected some shortages of context in the “*Map View*” and “*Timeline View*”. In detail, the simplifier proposed centering these views in “*location*” and “*time*” respectively. So, the simplifier proposed centering the “*Timeline View*” in the current time. Moreover, it took into account this context information centering the view in the next pending task. Similarly, the simplifier proposed centering the “*Map View*” on the user’s location. Thus, it proposed showing the current user’s location in order to inform where user is situated respecting to the pending tasks. Thus, user could plan in its mind an optimized route for completing its tasks.

Beyond showing more contextual information, some changes in the use of the context need to adapt the showed information. Several simplifier’s proposals represented this idea. In detail, one of its proposals was to adapt in the “*Timeline View*” the bar that informs about when the tasks must be performed according to the current date. So, as we see in the figure 4, it proposed showing the hour when a task must be performed in the system’s date and showing the date when a task must to be performed later.

Fig. 4. Changes in Timeline View after the simplifier's suggestions

Another example of how the simplifier proposed to adapt the showed information according the context is the proposal ofadapting how the markers were showed over the "*Map View*" according to the zoom level. In detail, the simplifier said that if a user had for example four tasks at the university, another in the museum and two more at home, the user could have totally seven tasks in its city. So, when user interacts with the zoom controls, if the map was focused over its city, it could represent three markers, but if the map was focused more generally, for example over Spain, these three markers could converge in just a marker which contains a seven and represents all the tasks of the city as we can see in the first image of figure 5. So, we contextualized the zoom controls while we reduce the amount of markers reducing also the number of components.

Fig. 5. Contextualized Map View

Add more intuitiveness. The main objective of providing intuitiveness is allowing users to interact with the applications naturally. Thus, taking little steps to more intuitive interactions could generate good feelings in users, so that we indirectly could generate emotions as the last guideline says.

So, the simplifier gave advices that moved popular knowledge to our particular case. One example about this idea is the represented in the figure 4. In this view, the simplifier proposed removing the horizontal scroll and the arrows that connect the tasks by taking benefit of gestures and of our natural association between the terms "*left*" with "*before*" and "*right*" with "*after*".Another example is how the simplifier avoided the use of the "*more details*" button in the overlay showed in figure 3. Instead of, the simplifier applied the concept of "*hyperlink*" in order to navigate to the more detailed view of the selected task as we see in figure 6.

Fig. 6. Simplified overlay for interacting with a task

As we have seen, the use of gestures provides more easy interactions. Another advice that provided intuitiveness in the "*Timeline View*" was allowing gestures to performing the "*complete task*" and "*delete task*". In detail, it proposed associating an upward movement over the representative icon of the task for changing the representative image of the task by a "*Completed*" message, and a downward movement for changing the representative icon of the task by a "*Deleted*" message. Similarly, it proposed a double tap in order to edit the task in the more detailed view of the task.

Save time. The goal of this point is to avoid user waste time because in the mobile context in which we center this work, time is appreciated. So, in our particular case, the simplifier proposed avoiding dialog alerts that were showed in the "*Timeline View*" when user deleted a task in order to confirm the operation. Instead of, the simplifier proposed that when user made a mistake and completed a task by means of a gesture, this action could be compensate rapidly by the use of another gesture.

Generate Emotions. The objective of generating emotions is providing users of good feelings in interactions in order to improve the satisfying. In our case, the simplifier didn't give any specific advice on this way, but, indirectly it provided good emotions when it proposed saving time or adapted information to the user's current context.

4 Conclusions

This work has provided an iterative process for the interaction design based on simplicity. We centered this process in mobile devices since the use of simple interactions fits perfectly the mobile device characteristics contributing to users focus on the task at hand. For this purpose, we added a new role centered in (1) finding complexities and (2) providing refactored solutions. For this purpose, we defined specific guidelines which adapt the Laws of Simplicity of Maeda in order to reach simplicity in interaction designs.

In our experience it wasn't needed a lot of time for applying the processin comparison with the provided benefits. In our study case, the simplifier intervened three times in interventions about 15 minutes long. Furthermore, we want remark that the designer was implicated in the process and usually accepted the simplifier's advices.Thus, we realized that when the process advanced, the designer collaborated in centering its view in the end user. However, thedesignercould bereluctant to make changes in defense of its designs and this could require more time in discussions. So, we want remark (1) the need of the simplifier's final validation and (2) theinfluence of the required time in the costs.

We found that designs were improved. But, prototyping is needed in order to include the users' opinionswith the purpose of verifyingthat the given advices don't remove value from the product and really reach simplicity. Thus, since the simplifier is not aware of the requirements, the designer must verify whether it is true or nor with the final user.

So, our next steps are (1) to evaluate our approach through obtaining user feedback from the prototype outside the lab and (2) to apply our approach in future prototype designs in which we can study new complexities which are provided by the interaction between users and its associated changes in the use of the context.

References

1. ISO DIS 9241-210:2008: Ergonomics of human system interaction - Part 210: Human-centered design for interactive systems (formerly known as13407). International Standardization Organization (ISO), Switzerland
2. Ju, W., Leifer, L.: The design of implicit interactions: Making interactive systems less obnoxious. Design Issues 24(3), 72–84 (2008)
3. Savio, N., Braiterman, J.: Design Sketch: The Context of Mobile Interaction. In: Proceedings of MobileHCI 2007, pp. 284–286 (2007)
4. Einstein, A.: Chapter 5: The Golden Rules of User Interface Design
5. Ladner, S.: The Essence of Interaction Design Research: A Call for Consistency. In: Interactions. Section: Challenges to Design Research table, vol. 17(2), pp. 48–51 (2010)
6. Maeda, J.: The Laws of Simplicity (Simplicity: Design, Technology, Business, Life). The MIT Press, Cambridge (August 2006)
7. Vastenburg, M.H., Keyson, D.V., de Ridder, H.: Considerate home notification systems: a field study of acceptability of notifications in the home. Personal and Ubiquitous Computing 12(8), 555–566 (2008)
8. Unger, R., Chandler, C.: A Project Guide to UX Design: For user experience designers in the field or in the making. New Riders Publishing, Thousand Oaks (2009)
9. Krug, S.: Don't make me think A Common Sense Approach to Web Usability
10. Villamor, C., Willis, D., Wroblewski, L.: Touch Gesture Reference (2010)
11. Cohen, M.H., Giangola, J.P., Balogh, J.: Voice: User Interface Design
12. Norman, D.A.: Human-centered design considered harmful. Interactions 12(4), 14–19 (2005)

Semantic Visualization of Wireless Sensor Networks for Elderly Monitoring

Carsten Stocklöw and Felix Kamieth

Fraunhofer Institute for Computer Graphics Research,
Fraunhoferstr. 5, 64283 Darmstadt, Germany
{carsten.stockloew,felix.kamieth}@igd.fraunhofer.de

Abstract. In the area of Ambient Intelligence, Wireless Sensor Networks are commonly used for user monitoring purposes like health monitoring and user localization. Existing work on visualization of wireless sensor networks focuses mainly on displaying individual nodes and logical, graph-based topologies. This way, the relation to the real-world deployment is lost. This paper presents a novel approach for visualization of wireless sensor networks and interaction with complex services on the nodes. The environment is realized as a 3D model, and multiple nodes, that are worn by a single individual, are grouped together to allow an intuitive interface for end users. We describe application examples and show that our approach allows easier access to network information and functionality by comparing it with existing solutions.

Keywords: Wireless Sensor Network, Semantic Visualization, Elderly Care.

1 Introduction

Wireless sensor networks in an AAL-context often serve the purpose of user monitoring like the collection of health data or detection of user location and movement. The currently dominant topology-based network visualizations mainly serve the purpose of network maintenance from a technology-focused point of view, highlighting network architecture features and connectivity. Visualizing the data in a *user-friendly fashion* requires a different approach to visualization. The network visualization is mapped to a 3D-model of its environment, locating the different nodes within the scene and changing the visualization based on the nodes' measured sensor data, thereby directly displaying the interpretation result *within* the visualization. This has distinct advantages in usability to the end-user, but also in network maintenance. By directly locating nodes in the environment and displaying the nodes' data in terms of interpreted results, malfunctions can be easily detected and located. Further visualization patterns, more in tune with the common logical approach to visualizing network topologies, can be layed over the fundament of a location-based 3D-visualization, thus also providing data on the nodes' connectivity in addition to location and functionality visualization.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 41–50, 2010.

We show that our approach is more flexible with respect to extendability, simulation, and debugging, and provides more practical insight into the monitored data than existing methods.

2 Related Work

Given the distributed deployment and the nature of wireless sensors, which typically do not have a display attached, an external visualization of the whole network, and on services as well as issues and problems of single nodes is essential for developers and for end users. Thus, a number of different visualizations have been proposed.

Most of the early visualization schemes focus on a logical illustration, e.g. showing the two-dimensional topology of the network using graph-based methods, like Spyglass [1]. Later visualizations show a model of the real world to allow a direct relation between the virtual and the real world. WIVA [6] uses realistic models for Structural Health Monitoring of buildings. Service data is shown as 3D objects, e.g. a fire sensor is shown as a 3D fire model. This way, it is not possible to integrate arbitrary service data without having a 3D representation of it. NetTopo [8] can be shown as 2D or 3D world, focusing on the logical topology of the network. In [5] a realistic world is shown, but no service data is displayed. The focus here is on finding the optimal placement. Also, [7] uses a real world visualization, i.e. virtual representations of sensor nodes layed over an image (e.g. a satellite image) of the area. However, no sensor data is displayed. Although these methods provide a real world visualization, only single and static, but no mobile nodes are considered.

In [2] one mobile node is used together with multiple static nodes to infer the location of the mobile node. This mobile node is directly connected to a PDA and can - together with a compass - show a virtual representation of the world from the users current position. Similar work is using multiple static nodes to calculate the position of a single mobile node, e.g. [4] for a 2D bird eye's view of the area. None of these methods considers groups of mobile nodes in a real world visualization.

3 System Overview

Over the years, many different types of WSNs have been developed for a variety of scenarios. Due to the restrictions of hardware capabilities, most WSNs are specialized. This section gives a coarse overview on what type of networks and nodes are taken under consideration for visualization.

3.1 System Architecture

The general system architecture is depicted in Fig. 1. Three different types of nodes are distinguished: static nodes, mobile nodes, and base station. *Static nodes* are single nodes, hidden in the environment. Their main purpose is to

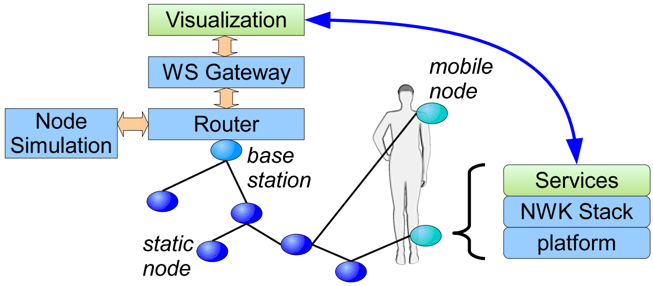


Fig. 1. System architecture

guarantee connectivity of all nodes. *Mobile nodes* are directly attached to an elderly, so that their position can change over time. Mobile as well as static nodes can have multiple sensors and multiple services running on them. Even complex services, like activity recognition, are possible. It is not known a priori which sensors and services are available on a specific node.

The last node type, the *base station*, connects the WSN to a software component, called router, on a PC. To allow the use of different WSNs at different locations there can be multiple routers that connect via IP or GPRS (for remote outdoor locations) to a single component that provides all functionality of the WSN, and acts as a central entrance point for front-end systems. The connection between this component - in this system called web service Gateway (WS Gateway) - and front-end is realized by a standardized interface, e.g. web services, but other interfaces are possible.

Through this abstraction of services the front-end system is independent from node hardware and communication protocols.

3.2 Simulation of Wireless Sensor Networks

Working with real nodes has some disadvantages for developers. Given the high budget requirements of sensor nodes, large deployments are sometimes not realized. Additionally, testing and debugging a simulated node on a PC is much more comfortable. Thus, simulation of nodes has gained increasing interest in WSN community. The system architecture (see Fig. 1) provides the possibility to add simulated nodes as well as real nodes; both are attached to a router and expose their services to the front-end using the same interface. An example for such an simulation environment is described in [3]. This way, the visualization can work with both real and simulated nodes in an integrated view.

3.3 Applications

The services running on the node are, in this system, uploaded to the node at any time. Thus, a different set of services can be present for a variety of use cases and end users.

A developer, for example, could use this method to show statistics of the networks routing layer, memory usage, or battery status. Furthermore, setting options is possible - like node type or node coordinates for static nodes - or calling functions - like rebooting a node. In the deployment and maintenance phase, information about network connectivity and reachability is important to guarantee reporting of alerts and to detect elderlies leaving a care center.

High-level classification of sensor data is essential for monitoring the health status of an individual. The temperature sensor can be used to detect a drop to very low levels causing *hypothermia*, *activity recognition* can be realized by an accelerometer sensor and show long periods of inactivity, and a cardiac ECG monitoring can detect heart rate irregularities (*arrhytmia*). Classification can generally be done on a node, but an alert should be triggered at a central point, e.g. a monitoring care center.

Localization of nodes is a fundamental application. It is necessary to correctly show the individuals in the virtual world and can be used by end users to detect the point of presence, e.g. for a medical emergency in a care center, this could lead medical experts directly to a patient.

These are just a few application examples. A visualization should be able to handle all use cases, provide a flexible and extensible framework, and an intuitive interface for the different end users.

4 Visualization

This section describes our novel method for visualizing groups of mobile nodes in a real world illustration.

4.1 Software Architecture

The general architecture of the visualization and its relationship to other components of the system is illustrated in Fig. 2(a). The connection to the WS Gateway is realized by a web service client to subscribe to services on the nodes, and by a web service server to receive information from the WS Gateway asynchronously. Both components, client and server, are used to create a virtual representation of the sensor network and provide methods to call node functions (called *actions*) and to subscribe to service output that is periodically and asynchronously delivered to an observer derived from the interface *interpreter*. In Fig. 2(b) a more detailed example for a temperature service is given.

The virtual representation also stores all information coming from nodes in a special data structure (Fig. 2(c)). For service data, static and mobile nodes are distinguished and for mobile nodes another layer is introduced to group nodes according to the individual wearing them. Access to this data is possible for modules and for services. Modules include functionalities for positioning of elderly models in the scene, showing different activities of an individual by changing the animation of that model, providing transformations between real world coordinates and virtual coordinates, and storing routing information to build up parent-child relationships of the routing tree, e.g. to show a connectivity graph. Services in the visualization framework often correspond directly to services on

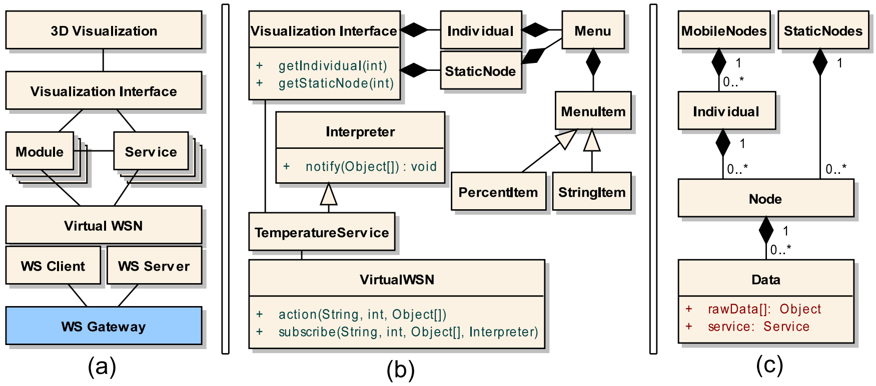


Fig. 2. Visualization architecture: (a) general architecture (b) detailed class diagram for temperature service (c) data model

the nodes. However, it is possible to create arbitrary services that do not need an on-node correspondence, but appear only in the visualization.

The 3D visualization is realized with VRML using an external viewer and the *external authoring interface* (EAI). To ease development of services and modules, an additional abstraction layer, the visualization interface, was implemented.

4.2 Real-World Visualization

The node network is functionally represented in a 3D real-world visualization. This visualization displays the rooms as well as the monitored elderly people in a way that maps the sensor node outputs to direct visual output in the 3D-scene.

Fig. 3 shows a screenshot of the visualization of an elderly care home with a set of elderly people in the scene. The basis of the visualization is the floorplan on which the positions of the different elderly people are being shown. Elderly people are represented with an animated 3D-model. This model changes based on the activity of the respective person as detected by the sensor nodes the person is wearing. Furthermore the visualization displays other functional system components like routers and static nodes (shown in Fig. 3 as a red ball in the bottom left). Another feature of the visualization is the integration of a *connectivity graph* into the visualization. This allows the user to see directly connections between the different nodes on different elderly people. This combination of localization of the nodes in space and their connections in the same visualization is a powerful tool for sensor network maintenance as it gives the engineer direct perception of which nodes have connection problems and which person is wearing it. This semantic visualization approach helps the end user by directly mapping the raw sensor data from the network to a model on human experience level which can be displayed directly in the visualization. Through this semantic approach a higher usability of the node network is being achieved.

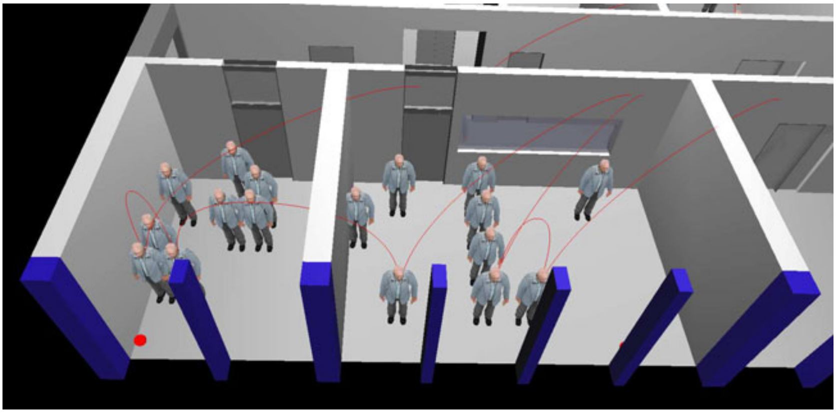


Fig. 3. Visualization of an elderly care home showing network connectivity

4.3 Localization for Elderly Model Placement

For a real-world visualization it is important to accurately place the 3D models of elderlies in the virtual environment. Three different localization methods were integrated based on the closest router, the closest static node and on real coordinates.

Localizing nodes using static entities can be realized on the routing layer. When a packet is routed from the sending node through several intermediate nodes, the first static entity on this route determines the approximate position of the sending node. The elderly wearing the sending node is then placed in a predefined area around the static entity. Accuracy strongly depends on the number and placement of static entities. It is also recommended, that static nodes are preferred over mobile nodes when creating the routing tree to make sure that the first static node forwarding a packet through the network is sufficiently close to the sending node. This way, it is possible to achieve room-level localization in case walls and ceilings are absorbing enough to prevent static nodes outside the room of presence to act as parent in the routing tree.

The third localization method is based on real coordinates. This can be achieved by measuring the *received signal strength indication* (RSSI) of surrounding nodes with known position and using lateration for position estimation. Taken into account the measurement errors for RSSI values, a more accurate position estimation can be achieved if groups instead of single mobile nodes are observed.

Additionally, an area in the virtual world can be defined for nodes of unknown position. If the location could not be determined, the elderly 3D model is placed in this area. This can also be used to directly show the absence of persons in an elderly care center or to indicate a connectivity problem.

4.4 Visualization of Sensor Data

The visualization provides access to the information from the nodes in the network by displaying it in a list of menus, which can be activated by clicking on one of the models of elderly people. These menus list detailed information about the measurements performed by the nodes worn by an elderly person as well as options for interacting with the nodes. Each menu provides information on a single node. The user can navigate through the list of menus through the use of the buttons at the bottom of the list. The menu offers three different kinds of items: *Text items*, which display information on the node in the form of text. *Labeled number items*, which display numerical data from the node with a description of what is being displayed. *Labeled button items*, which let the user interact with the node by clicking on the displayed button, the function of which is being described in a label.

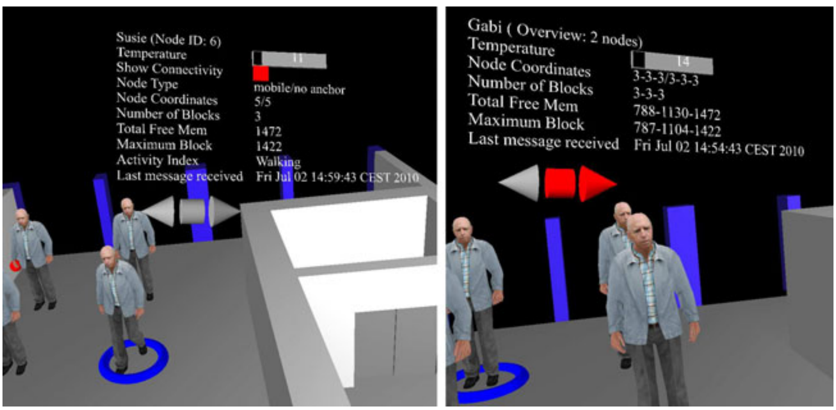


Fig. 4. Sensor data for single node (left) and overview over multiple nodes (right) showing minimum, average, and maximum of all nodes attached to an individual

In addition to menus on individual nodes, for each elderly person there is an *overview menu*, which provides aggregated information about all the nodes worn by the person. Fig. 4 shows both an individual node menu as well as an overview menu for an elderly person. The menu on the left hand side shows the different information available on a single node. First the name of the person wearing the nodes is given followed by the unique identifier of the displayed node. The next item displays the temperature value measured by the node. This is followed by a labeled button item, which lets the user display the connectivity graph for this node in the visualization. The next items describe further technical information on the nodes. The arrows at the bottom let the user navigate to the next node on the elderly person, while the button in the middle makes the visualization display the overview menu. The arrows are functional if they are colored and non-functional if they are grayed out - for example when there are no nodes with a higher ID the right button is deactivated. The image on the right shows the overview menu of an elderly person. The temperature value provides the average value of all the nodes on that person. Measured coordinate values and memory information are given for all nodes respectively. The last message received gives the last message sent by any of the nodes. This should serve as an example of the different services which can be connected to the visualization. One important feature of the visualization is the ability to connect to any kind of service and make it available through one of the menu item templates in the node menu.

4.5 Extendability

As new services can be integrated into the system at any time, it is necessary to provide an extensible and flexible framework for the visualization as well. Thus, both sides, the wireless sensor network as back-end and the display as front-end, are abstracted for services, so it is not necessary to handle connection issues or use low-level VRML constructs. To create a new plugin, a developer has do derive from a specific class, subscribe to a service to get the appropriate data, and override a method to draw the menu with service data according to the menu templates provided. Additionally, the visualization interface provides methods to modify the virtual world in a variety of ways, e.g. show/hide individuals, or set individual position and status.

5 Comparison with Existing Solutions

The work presented in this paper is compared with existing solutions in the following aspects:

Extendability. Most of the related work do not provide an interface for new services. Spyglass [1] provides an API, but uses a two-dimensional display. The focus here is on providing a three-layer concept for gradient information (e.g. a temperature gradient), relations between nodes (e.g. routing tree) and node information (e.g. absolute sensor data). As we are using a real world illustration in 3D, a layered concept is not needed. SNAMP [9] uses a different approach by introducing different views. This way, it may not be possible to see different information without changing to a another view, whereas we provide an integrated method.

Simulation and Debugging. As described in section 3.2 it is possible to add simulated nodes to the network and work with both, real and simulated nodes, in an integrated view. Having a complex node simulator, even debugging of node programming code can be achieved. Additionally, special services can be uploaded to a node to show node properties, e.g. memory information or routing statistics. Existing solutions mainly focus on one of the aspects simulation or visualization. NetTopo [8] provides such a mechanism, but the visualization is restricted to a logical (e.g. graph topology) representation.

Scalability. Visualizing objects with complex 3D models allows an intuitive interface but at the expense of scalability. Although the maximum number of individuals that can be shown strongly depends on the hardware capabilities and the models complexity, the number of nodes is not that much affected. If multiple nodes are attached to an elderly, only one node is shown.

6 Evaluation

The graphical user interface has been evaluated with nine participants using a questionnaire based on the ISO9241-110 standard. Each participant received a set of tasks to perform with the developed tool and was then asked to complete the questionnaire and give a short written commentary to the software. The evaluated criteria are illustrated in Fig. 5. There are two additional criteria in the standard, *error tolerance* and *suitability for individualization* - they have not been evaluated, since the questions used in the ISO9241-110 evaluation standard do not target this GUI concept implementation and have been subsequently left out. The evaluation results have been promising.

Fig. 5. Evaluation results

Fig. 5 shows the evaluation analysis with average values for test results and standard deviation intervals for all questions asked. The test results can range from 1 to 7, 1 meaning the worst evaluation, 4 being neutral and 7 being very positive. The range of all results is in the upper hand spectrum above a value of 4, meaning overall positive test results. As can be seen in the graph, most results are in the range of six, the second best test result, with only self-descriptiveness at an average of slightly above 5. From the commentaries to the evaluation it could be extracted that while the overall impression was clearly positive (as indicated in the evaluation results), there were some things that were not as intuitive as initially thought. First, the names given to the different fields in the 2D-part of the GUI left some participants undecided about which field to pick to choose a person in the simulation, for example. Other suggested improvements are a direct scrolling to chosen nodes and an instant activation of the node menu in the 3D-visualization when a person or a node has been chosen in the 2D-GUI.

7 Conclusion

This paper presented a novel approach to the visualization of wireless sensor networks focusing on a realistic representation of the nodes' environment combined with dynamic access to the services provided on the network nodes as well as a visualization of node connectivity and state. This semantic visualization provides the end user a with a direct visual access to network information which simplifies both elderly monitoring and node network maintenance. It abstracts from raw sensor data and maps it to a model on the level of human knowledge, representing the information in direct visual form. The use of menu templates allows the user to utilize nodes with a variety of different offered services and have them available in the graphical user interface. It has also been outlined how this work distinguishes itself from and improves upon prior work in the three main areas of extendability, simulation and debugging, and scalability. Additionally, an evaluation was performed to validate the effectiveness of our approach.

Acknowledgements. This work is partially financed by the European Commission under the FP6 IST Project *Wirelessly Accessible Sensor Populations (WASP)*, contract IST-034963.

References

1. Buschmann, C., Pfisterer, D., Fischer, S., Fekete, S.P., Kröller, A.: Spyglass: a wireless sensor network visualizer. SIGBED Rev. 2(1), 1–6 (2005)
2. Chung, W.Y., Lee, B.G., Yang, C.S.: 3d virtual viewer on mobile device for wireless sensor network-based rssi indoor tracking system. Sensors and Actuators B: Chemical 140(1), 35–42 (2009)
3. Fraboulet, A., Chelius, G., Fleury, E.: Worldsens: Development and prototyping tools for application specific wireless sensors networks, pp. 176–185 (April 2007)
4. Klingbeil, L., Wark, T.: Demonstration of a wireless sensor network for real-time indoor localisation and motion monitoring. In: International Conference on Information Processing in Sensor Networks, IPSN 2008, pp. 543–544 (2008)
5. Kockara, S., Dagtas, S.: Towards 3-dimensional and optimized design of wireless networks for smart environments. J. Comput. Small Coll. 21(5), 145–153 (2006)
6. Koo, B., Choi, H., Shon, T.: Wiva: Wsn monitoring framework based on 3d visualization and augmented reality in mobile devices. In: Mokhtari, M., Khalil, I., Bauchet, J., Zhang, D., Nugent, C. (eds.) ICOST 2009. LNCS, vol. 5597, pp. 158–165. Springer, Heidelberg (2009)
7. Miyashita, M., Nesterenko, M., Shah, R., Vora, A.: Visualizing wireless sensor networks: An experience report. In: Yang, L.T., Arabnia, H.R., Wang, L.C. (eds.) ICWN, pp. 412–419. CSREA Press (2005)
8. Shu, L., Wu, C., Zhang, Y., Chen, J., Wang, L., Hauswirth, M.: Nettopo: beyond simulator and visualizer for wireless sensor networks. SIGBED Rev. 5(3), 1–8 (2008)
9. Yang, Y., Xia, P., Huang, L., Zhou, Q., Xu, Y., Li, X.: Snamp: A multi-sniffer and multi-view visualization platform for wireless sensor networks. In: 1st IEEE Conference on Industrial Electronics and Applications, 2006, pp. 1–4 (2006)

Privacy Management and Control in ATRACO

Bastian Könings, Björn Wiedersheim, and Michael Weber

Institute of Media Informatics, Ulm University, Germany
{bastian.koenings,bjoern.wiedersheim,michael.weber}@uni-ulm.de

Abstract. Privacy is a critical factor for the acceptance and success of next generation ambient intelligent environments. Those environments often act autonomously to support a user's activity based on context information gathered from ubiquitous sensors. The autonomous nature and their accessibility to large amount of personal information raises several privacy issues for participants in such environments. Those issues need to be addressed by adequate privacy mechanisms. In this paper we present an overview of the ATRACO approach to provide privacy management and control in an ambient intelligent environment.

Keywords: privacy, ambient intelligence, ubiquitous computing.

1 Introduction

In the last decades several research has been done in the field of *Ambient Intelligence* (AmI), or Ubiquitous Computing, which offer great opportunities for a large number of new applications ranging from assisted or daily living systems, entertainment systems, to intelligent transportation systems. Such systems will be invisibly embedded into our everyday environments through a pervasive transparent infrastructure consisting of a multitude of sensors, actuators, processors and networks. The interplay of those components allows the system to support, interact with and adapt to individuals in a seamless and unobtrusive way.

However, in order to allow such a high flexibility, support and adaptation, AmI systems require a large amount of information, such as real-time information gathered from ubiquitous sensors, personal user information, and the ability to intervene in the user's physical environment. These requirements raise several privacy issues which need to be addressed in order to satisfy user needs and acceptance.

The ATRACO (Adaptive and TRusted Ambient eCOlogies) project [9] aims to realize an adaptive AmI system based on the concept of *Activity Spheres* (AS). An AS is the utilization of knowledge, services and other resources required to realize an individual user goal within an AmI environment. This paper presents the privacy concepts embedded in the overall ATRACO architecture.

The main ATRACO components are described in Section 2. Section 3 provides and overview of related work in the area of privacy in ubiquitous computing. An overview of the general privacy concepts in ATRACO is given in Section 4 followed by a detailed presentation of the ATRACO privacy components in Section 5. The paper is concluded by Section 6.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 51–60, 2010.

2 ATRACO Architecture and Components

The ATRACO [9] approach aims to realize the concepts of an AmI system by addressing heterogeneity of artifacts, system transparency, discovery & management of various artifacts, and autonomous behavior of learning agents. ATRACO uses a Service-Oriented Architecture (SOA) at the resource level to support numerous devices and sensors, and at the system level to support ubiquitous computing applications. Several agents complement the SOA infrastructure by providing high level adaptation to a user's task. ATRACO agents support adaptive planning, task realization and enhanced human computer interaction. The concept of ontologies is used to address the semantic heterogeneity that arises in AmI environments. Ontologies provide a common repository of system knowledge, policies and state. By combining the above approaches, ATRACO achieves a high degree of adaptation.

The main components of the ATRACO architecture are the *Sphere Manager*, the *Ontology Manager*, the *Fuzzy Task Agent*, the *Planning Agent*, and the *Interaction Agent* as depicted in Figure 1. We will briefly describe the functionality of each component in the following.

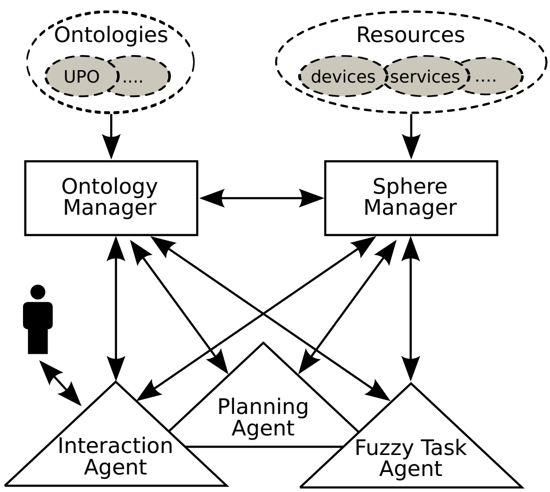


Fig. 1. Main components of the ATRACO architecture

Sphere Manager. The Sphere Manager (SM) is responsible for initializing or dissolving an Activity Sphere (AS) with its associated resources for a specific user goal. The SM acts as an event service to other components of the AS. It monitors the execution of the task workflow and adapts the composition of resources in case of any conflicts, such as failing devices or other exceptions.

Ontology Manager and Ontologies. Ontologies provide a central knowledge base for properties, state and context information as well as user preferences. The main sphere ontology is formed by merging or aligning of more specific domain ontologies, task ontologies and the *Upper Level Ontology* (ULO), which describes basic entities and relationships of the ATRACO world model. The most privacy relevant ontology is the *User Profile Ontology* (UPO), which contains user preferences, user properties and other personal related information. The Ontology Manager (OM) is responsible for managing the sphere ontology and responds to ontology queries of other components.

Fuzzy Task Agent. To deal with the huge amount of environmental and user-based uncertainties the Fuzzy Task Agent (FTA) learns and adapts its rule set based on general type-2 fuzzy logic. This enables facing changes in the environment, in user behavior and user preferences.

Planning Agent. The Planning Agent (PA) participates in the realization and adaptation of activity spheres. The PA encapsulates a search engine that exploits hierarchical planning and partial-order causal-link planning. It delivers workflows consisting of basic executable tasks needed to achieve a user's goal.

Interaction Agent. The Interaction Agent (IA) offers multimodal services to interact with the user and provides an interface to the system. In order to provide smooth interaction the IA uses distributed multimodal interaction widgets which adapt to the user's context. Therefore, the user interaction can be distributed among available modalities and devices at runtime.

3 Related Work

Already Marc Weiser, who shaped the vision of ubiquitous computing [17], recognized that AmI environments will raise new privacy issues. However, his privacy concerns were focused on location privacy. Location privacy is a particular form of information privacy and has attracted considerable research [2, 10].

In recent years, general information privacy in ubiquitous computing has been investigated as well. Hong and Landay [7] proposed Confab, a privacy-aware architecture for ubiquitous computing. Confab uses the concept of information spaces by Jiang et al. [8] in combination with in- and out-filters to manage the flow of context information about a person. Langheinrich [12] discussed how the fair information practices, a collection of abstract guidelines to achieve information privacy, can be adapted to ubiquitous computing scenarios. He pointed out six principles, which he followed to design the privacy awareness system PawS [13] in order to address third party data collection in smart environments.

Privacy protection mechanisms, or *Privacy Enhancing Technologies* (PETs), can in general be divided into the three main categories of *policy matching*, *prevention & control* and *detection*.

Policy matching techniques try to minimize the privacy risks by specifying and matching user privacy policies with involved service provider policies. The most common policy languages of that kind are the *Platform for Privacy Preferences* (P3P) [5], the *Enterprise Privacy Authorization Language* (EPAL) [1] and the *eXtensible Access Control Markup Language* (XACML) [6], which are XML-based languages to support the definition and enforcement of privacy policies and obligations. The main drawback of privacy policy systems is that they generally cannot enforce privacy and instead rely on trustworthiness and regulatory pressures to ensure policy compliant behavior.

Prevention and control mechanisms seek to prevent misuse of personal information by adopting information in several dimensions or by avoiding privacy-critical operations or access. Such mechanisms are *Pseudonymization* [4], *Obfuscation* [19] or *Access Control* [14].

Detection mechanisms try to identify and penalize privacy violators. A system for privacy violation detection, such as PRIVDAM [3], continuously monitors access to personal data and detects misuse or abnormal behavior.

4 Privacy in ATRACO

Privacy in the context of ATRACO comprises two forms of privacy, namely *information privacy* and *territorial privacy*. Information privacy refers to the protection of personal information (or personal data) whereas territorial privacy refers to the protection of private spaces (or territories). This classification will be discussed in detail in Section 4.1. The basic approach to enforce privacy in ATRACO is based on policy matching techniques and access control mechanisms. All privacy relevant functionalities and components are encapsulated within the *Privacy Manager* (PM). The integration of the PM within the overall architecture and its interaction with the Sphere Manager (SM), Ontology Manager (OM) and Interaction Agents (IA) is discussed in Section 5.

4.1 Privacy Classification

To understand how privacy can be protected in the context of an AmI Environment, we first need to clarify how the term *privacy* is represented in ATRACO. Formulating an adequate definition of privacy is one of the most intractable problems in privacy research. Besides hundreds of existing definitions one of the oldest still remains influential: Samuel Warren and Louis Brandeis's 1890 declaration of privacy as the *"right of an individual to be let alone"* [16]. We refer to this more traditional privacy point of view as *Territorial Privacy*. Later in 1967 Westin defined privacy as *"the claim of individuals, groups or institutions to determine for themselves when, how, and to what extent information about them is communicated to others"* [18]. Many of today's common definitions and understandings of privacy in the context of information technology systems are based on the latter definition and are referred to as *Information Privacy*. Both privacy categories are prevelant in ATRACO:

Information Privacy. In ATRACO, information privacy aims to protect any kind of sensitive information from leaving the borders of an activity sphere. Personal information must remain at any time under full control of the user. Personal information includes *personal data* (e.g., pictures, e-mails, web history), *personal properties* (e.g., age, size, weight), *personal preferences* (e.g., favourite music, movies, fashion) or *personal behavior* (e.g. doing what, when, how often).

Personal information can either be static (e.g., mails, pictures) or dynamic (e.g., location, activity) and can be collected directly (e.g., database access, physically sensed) or indirectly (e.g., analytically derived, compositionally derived). The main knowledge base for such personal information is the User Profile Ontology (UPO). Thus access control to the UPO is a main design goal of the ATRACO privacy components. However, sensitive information in other ontologies is protected as well.

Territorial Privacy. While protecting information privacy is sufficient in the context of most common IT domains, it is only one aspect of privacy protection in the context of AmI systems. The pervasive nature of such systems equipped with multiple sensors, actuators and computational devices, constitute a new facet of privacy protection, that we call *territorial privacy.* Territorial privacy aims to satisfy a more traditional expectation of privacy such as being in *"a state in which one is not observed or disturbed by others"* [15].

As all users present in an AmI system might be continuously observed by cameras, microphones or other sensors, it is required that a user is able to control whether or not an entity is allowed to observe the user in a certain way. We call the different ways of observation *observation channels.* For example, an entity receiving a life video stream of a user is connected by a visual observation channel. A detailed discussion and formalization of this observation model is given in [11]. Further, the control of how other entities are allowed to disturb a user in his private space is an important aspect of territorial privacy. The level of disturbance can be either physical, e.g., if a person enters the room, or virtual. Virtual disturbances, for instance, are undesired outputs of visual or acoustic signals or undesired initiations of interactions.

4.2 Privacy Requirements

For the design of the ATRACO privacy components we identified the following privacy requirements:

1. Privacy protection must depend on user privacy preferences and context.
2. User privacy preferences must be represented as privacy policy ontologies.
3. Privacy protection must be deployed at a sphere level when sensitive data is requested, or when the user wants to reduce observations or disturbances.
4. A user should be aware of any event relating to privacy.

4.3 Privacy Policy Ontologies

In general a *Privacy Policy Ontology* (PPO) describes how, under which conditions, and in which context an entity is allowed to handle personal information or is allowed to participate in a user's activity. In case of information privacy, policies need to be specified on the user side as privacy preferences and on the receiver side as privacy obligations in order to allow a matching between the two parties. The policies of receivers (e.g., persons, web services, but also system components) can either be specified as policies in PPO representation or in a different policy language, like P3P. For the latter case a policy wrapper translates common policy languages into PPO representation.

We distinguish between *Information Privacy Policies* (IPPs) and *Territorial Privacy Policies* (TPPs). The key attributes of an IPP are *information item* (the information to apply the policy to), *purpose* (what is allowed to be done with the information), *recipient* (who is allowed to use the information), *retention* (how long the information can be used), and *context* (further access constraints).

A TPP consists of the key attributes *territory* (the space to apply the policy to, explicitly identified by a location, or implicitly by an activity), *participant* (the entity who is allowed to observe or disturb the user), *observation type* (the allowed observation channel, declared by sensor types), *disturbance type* (the allowed type of intervention), and *context* (further access constraints).

In our prototype implementation we realized policies as part of the User Profile Ontology (UPO). Each policy is an individual of one of the two class types `InformationPrivacyPolicy` or `TerritorialPrivacyPolicy`. Policies of the first type may permit access to specific private information, indicated by the object property `canUse`, to different entities via the `permitsAccess` object property. If access to private information should be granted only for specific actions, a policy can specify the `allowsAction` object property, accordingly. Further, time constraints can be specified with the data properties `hasStartDate` and `hasEndDate` to limit the usage of private information to a predefined time interval. Figure 2 shows a simplified example of a IPP, granting access to view a picture for a time period of one year to Bob.

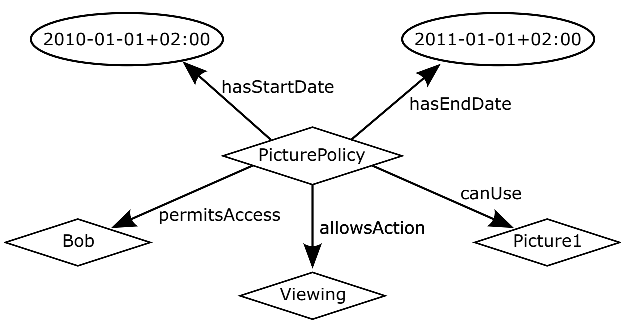


Fig. 2. Example of a Privacy Policy Ontology

5 ATRACO Privacy Components

In ATRACO, all privacy components are encapsulated in the Privacy Manager (PM) consisting of the *Policy Processing Engine* (PPE), the *Information Privacy Controller* (IPC), the *Territorial Privacy Controller* (TPC) and the *Feedback & Control Component* (FCC). Figure 3 provides an overview of the main components of the ATRACO privacy architecture and its interactions with the Sphere Manager (SM), Ontology Manager (OM) and Interaction Agent (IA).

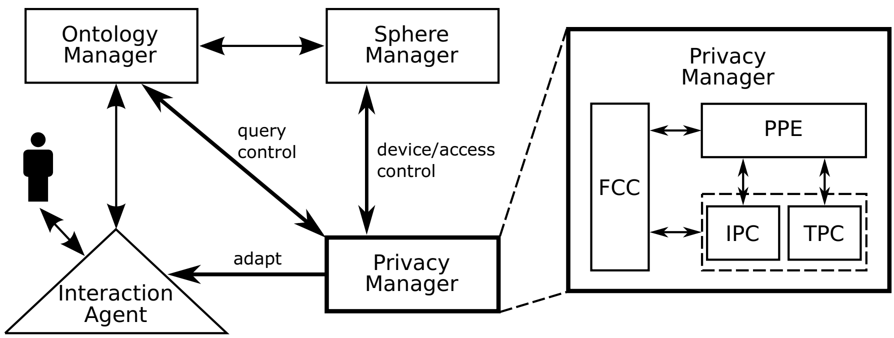


Fig. 3. Integration of the Privacy Manager

5.1 Privacy Manager

The Privacy Manager (PM) serves as a privacy interface for other components in the ATRACO architecture. Whenever a privacy relevant event occurs, that is a query of personal information, territory access requests, or changing context, the PM requests related policies and contextual information from the OM. The Policy Processing Engine (PPE) uses this information to perform policy reasoning and provides the results to the corresponding privacy controller, which will perform further privacy enhancing mechanisms based on the results. If any conflict occurs during this reasoning process the PPE can trigger the IA via the FCC in order to ask the user for a conflict resolution, such as modifying policies or adding policy exceptions. Further, the FCC provides a user interface for controlling and modifying privacy settings. The interplay and role of privacy controllers is described together with the discussion of potential privacy affecting events:

Queries of Personal Information. To ensure information privacy the OM triggers the PM to validate queries on ontologies, in particular on the UPO. The privacy validation will be handled by the PPE, involving the requester ID, associated policies, and current context. Based on the validation result the Information Privacy Controller (IPC) will grant or deny access to the query results.

Optionally, the IPC can perform obfuscation and pseudonymization techniques to further enhance privacy or react to policy conflicts.

A special case occurs whenever IAs request personal information. The IPC then checks if any other persons are located inside the user's sphere by querying the OM. If this is the case, the IPC uses the IA's ontology to determine the modality the information is going to be presented on and whether it is accessible by other persons. If the modality is accessible by other persons, for example a large screen, the PPE checks the related policies as described before. On a deny decision, the IPC may either retain the personal information or trigger the IA to adapt the interaction modalities to ensure that unprivileged access to personal information is prevented.

Territory Access Requests. The Territorial Privacy Controller (TPC) provides the functionality of territorial access control, in order to protect territorial privacy whenever external entities attempt to join the user's sphere. Access requests are forwarded by the SM. An access request may either ask for an observation or disturbance. Observation access is requested when an entity attempts to receive sensed user data, for example via a camera or microphone. A disturbance access is requested whenever an entity attempts to pass the physical borders (e.g. enter a room) or wants to actively intervene in the user's territory, for example, by acoustic or visual outputs, or by initiating interactions. Access decisions are based on the associated policies, which are validated by the PPE.

Contextual Changes. Contextual changes, such as changing activity, location or newly appearing persons, are reflected as events by the SM. Whenever such an event occurs, the IPC and TPC perform appropriate privacy measures. In case of an appearing person the IPC checks, whether personal information is currently presented by any interaction modality the appearing person would have access to. This information is either obtained from the interaction ontology or from the interaction agent directly. If so, the IPC may again adapt the interaction modalities based on the policy reasoning results of the PPE, as described before. In case of changes in activity or location, the TPC checks if any device or remote entity is currently observing or disturbing the user in his new territory. This process is analogous to the IPC's operation. The PPE validates the associated privacy policies. Based on the results, the TPC will exclude undesired observing and disturbing entities by switching off relevant sensors or active devices, respectively.

5.2 Prototype Evaluation

A prototype of the overall ATRACO architecture has been deployed in the iSpace at the University of Essex in order to perform evaluations in a real world testbed. A first evaluation of the ATRACO privacy components has been performed as part of a user test within the project's second year time period. We will briefly describe the privacy relevant scenario and results in the following.

Scenario. A participant, Alice, sits on the sofa at home. The system asks Alice if she wants to look at holiday pictures on the TV. Alice accepts and the slideshow starts. Some of the pictures have been marked as *private*, i.e. only accessible to the owner. After a couple of minutes a guest arrives at the front door. The system informs Alice that "*Bob has arrived*" and asks "*would you like him to enter?*". The participant responds by saying "*Yes*". The door opens and as Bob enters the private pictures are seamlessly removed from the slideshow.

Results. The idea to automatically hide private information in presence of other persons was in generally accepted by the participants. However, participants were more concerned about the ability to control the system. In situations where the user felt out of control, reactions to the system were mostly negative. In some cases, participants were comfortable with the system initiating interactions, such as when a guest arrives at the door, while in other cases, it was regarded as an invasion of personal space, for example, when participants were asked if they would like to look at some pictures. Perceptions of control may also be influenced by the channel of interaction. Voice interactions appeared to imply the existence of a separate social presence which could lead to a feeling that the personal space has been invaded. Also mood changes could have an impact on the way a user would prefer to interact with the system.

The results verify that privacy concerns of participants in AmI environments often refer to territorial privacy aspects, such as the fear of losing control and the perception of privacy invasions or particular forms of interaction. However, these aspects are already covered by our privacy architecture and will be further integrated in interaction behavior of the prototype for the next evaluation phase.

6 Conclusion

We have proposed the conceptual and architectural approach for managing and controlling information privacy and territorial privacy in ATRACO, an Ambient Intelligence (AmI) environment. While information privacy has been the dominating aspect of privacy research in recent years, next generation AmI environments raise the demand for more traditional privacy aspects, referred to as territorial privacy. Our presented privacy components were designed to achieve privacy by respecting user privacy preferences and contextual information. The integration of these privacy components in the overall ATRACO architecture presents a first step towards a trustworthy AmI system. However, a first evaluation has shown that further steps are necessary to satisfy user needs and acceptance. One next step will be the design and integration of a trust managing component, which will further improve privacy decisions and adaptation.

Acknowledgments. The research leading to these results has received funding from the European Community's 7^{th} Framework Programme (FP7/2007-2013) under grant agreement No. 216837 (project ATRACO).

References

[1] Ashley, P., Hada, S., Karjoth, G., Powers, C., Schunter, M.: Enterprise privacy authorization language (EPAL 1.2). Tech. rep., W3C (2003), http://www.w3.org/Submission/2003/SUBM-EPAL-20031110/
[2] Beresford, A., Stajano, F.: Location privacy in pervasive computing. Pervasive Computing, IEEE 2(1), 46–55 (2003)
[3] Bhattacharya, J., Dass, R., Kapoor, V., Gupta, S.: Utilizing network features for privacy violation detection. In: Proc. of the 1st Intl. Conference on Communication System Software and Middleware, pp. 1–10 (2006)
[4] Chaum, D.L.: Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM 24(2), 84–90 (1981)
[5] Cranor, L., Dobbs, B., Egelman, S., Hogben, G., Humphrey, J., Langheinrich, M.: The platform for privacy preferences 1.1 (P3P1.1) specification. Tech. rep., W3C (2006), http://www.w3.org/TR/2006/NOTE-P3P11-20061113/
[6] Godik, S., Anderson, A., Parducci, B., Humenn, P., Vajjhala, S.: OASIS eXtensible access control 2 markup language (XACML) 3. Tech. rep., OASIS (2002)
[7] Hong, J.I., Landay, J.A.: An architecture for privacy-sensitive ubiquitous computing. In: Proc. of the 2nd Intl. Conference on Mobile Systems, Applications, and Services, pp. 177–189. ACM, Boston (2004)
[8] Jiang, X., Hong, J., Landay, J.: Approximate information flows: Socially-Based modeling of privacy in ubiquitous computing. In: Borriello, G., Holmquist, L.E. (eds.) UbiComp 2002. LNCS, vol. 2498, pp. 176–193. Springer, Heidelberg (2002)
[9] Kameas, A.D., Goumopoulos, C., Hagras, H., Callaghan, V., Heinroth, T., Weber, M.: An architecture that supports Task-Centered adaptation in intelligent environments. In: Advanced Intelligent Environments, Boston, USA. Lecture Notes in Electrical Engineering, pp. 41–66 (2009)
[10] Krumm, J.: A survey of computational location privacy. Personal and Ubiquitous Computing 13(6), 391–399 (2008)
[11] Könings, B., Schaub, F., Kargl, F., Weber, M.: Towards territorial privacy in smart environments. In: Intelligent Information Privacy Management Symposium (Privacy 2010). Stanford University, USA (2010)
[12] Langheinrich, M.: Privacy by design - principles of Privacy-Aware ubiquitous systems. In: Abowd, G.D., Brumitt, B., Shafer, S. (eds.) UbiComp 2001. LNCS, vol. 2201, pp. 273–291. Springer, Heidelberg (2001)
[13] Langheinrich, M.: A privacy awareness system for ubiquitous computing environments. In: Borriello, G., Holmquist, L.E. (eds.) UbiComp 2002. LNCS, vol. 2498, pp. 237–245. Springer, Heidelberg (2002)
[14] Ni, Q., Bertino, E., Lobo, J., Calo, S.B.: Privacy-Aware Role-Based access control. Security & Privacy, IEEE 7(4), 35–43 (2009)
[15] "Privacy" Definition: The Oxford English Dictionary, 2nd edn. Oxford University Press, USA (2005)
[16] Warren, S., Brandeis, L.: Right to privacy. Harvard Law Review 4, 193–220 (1890)
[17] Weiser, M.: Some computer science issues in ubiquitous computing. Communications of the ACM 36(7), 75–84 (1993)
[18] Westin, A.F.: Privacy and freedom. Atheneum, NY (1967)
[19] Wishart, R., Henricksen, K., Indulska, J.: Context privacy and obfuscation supported by dynamic context source discovery and processing in a context management system. In: Indulska, J., Ma, J., Yang, L.T., Ungerer, T., Cao, J. (eds.) UIC 2007. LNCS, vol. 4611, pp. 929–940. Springer, Heidelberg (2007)

Place in Perspective: Extracting Online Information about Points of Interest

Ana O. Alves[1,2], Francisco C. Pereira[1], Filipe Rodrigues[1], and João Oliveirinha[1]

[1] CISUC, University of Coimbra, Portugal
{ana,camara}@dei.uc.pt, {fmpr,jmforte}@student.dei.uc.pt
[2] ISEC, Coimbra Institute of Engineering, Portugal
aalves@isec.pt

Abstract. During the last few years, the amount of online descriptive information about places has reached reasonable dimensions for many cities in the world. Being such information mostly in Natural Language text, Information Extraction techniques are needed for obtaining the *meaning of places* that underlies these massive amounts of commonsense and user made sources. In this article, we show how we automatically label places using Information Extraction techniques applied to online resources such as Wikipedia, Yellow Pages and Yahoo!.

1 Introduction

In this paper, we present our approach to the challenge of assigning semantic annotations to places. These annotations are automatically extracted by applying web mining and information extraction techniques that have been thoroughly applied and tested in previous works[1]. In our case, we are particularly focused on extracting information that allows an interpreter to distinguish a place from other places that are spatially or conceptually close. In other words, the *meaning of a place* is a function of its most salient features, present in the textual descriptions found in online resources about that place. In our case, places correspond to *Points Of Interest* (POIs), as these are abundant in the web. By definition, a POI is a place with meaning to someone and, if it is available online, it is likely that its interest is shared by many people. In our approach, we first crawl the web to get a large quantity of POIs and then analyze each of them in order to obtain their individual *semantic index*: the set of words that best define it.

A system that is able to extract relevant semantics from places can be useful for any context aware system that behaves according to position. The level of information considered in this paper brings another layer to add to other sensors (GPS, accelerometer, compass, communications, etc.), eventually pushing forward the potential for intelligent behaviour. We believe that Location Based Services can improve their perception of Location Context through Semantic Enriched Places. Being able to infer implicit properties about places by semantic tags, they may relate places semantically closed that by classical representation

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 61–72, 2010.

(position, name, category) it would not be so cleary indentified. For example, a machine learning algorithm in a smartphone could be trained to present a different interface according to type of place (e.g. leisure, work, shopping). Other uses can be imagined, from navigation applications (e.g. navigating by concepts, searching for a place given related words) to analysis of social interactions and space use (e.g. finding correlations between POIs and presence of people). Beyond the scope of this article, we have also explored the dynamic information related to the places analysing the events happening in a given city.

We will start by giving the reader some essential background, and then we explain our methodology. We also present experiments and discuss a validation of the results. The present work is expected to be made publicly available by the expected date of publication of the article.

2 State of the Art

2.1 Semantics of Place – From Space to Place

The difficulty in the unambiguous conceptualization of place comes along with its association to *space* and with the amount of different f that may arise. Consider the simple question "Where am I?" and a sample of possible answers: relative to function ("I'm at work"); relative to someone ("I'm at my friend's place", "I'm with John"); relative to scale ("I'm in the US", "I'm in New York"; "I'm in 14th Street"); relative to objects ("I'm in my car", "I'm outside the stadium"). To this list of physical references, we can add the wealth of metaphoric creations of place (e.g. "I'm in second life", "My mind was somewhere else"). A place can be described with geographic, demographic, environmental, historical, and, perhaps also commercial attributes. The meaning of place derives from social conventions, their private or public nature, possibilities for communication, and many more [2,3]. Perhaps a simplification in this context, Harrison [4] works on the distinction between the concept of place from space, a place is generally a space with something added - social meaning, conventions, cultural understandings about role, function and nature. Often, it also has temporal properties; the same space can become different places at different times. Thus, a place exists once it has meaning for someone and the perception of this meaning is the main objective of our research.

As pointed by [5], absolute position such as the pair latitude/longitude is a poor representation of place. In our point of view, flexible representations that allow different perspectives become of greater importance, describing the world by commonsense and human-recognizable labels that best illustrate, in a synthetic way, the distinctive features of a given place contained in it (be it just a Point of Interest or a broad geographic Area).

2.2 Automatic Tagging – From Text to Terms

Lemmens and Deng [6] proposed a semi-automatic process of tag assignment which integrates knowledge from Semantic Web ontologies and the collection of Web2.0 tags. On a different direction, Rattenbury et al [7] identify places and events from tags that are assigned to photos on Flickr. They exploit the regularities on tags in which regards to time and space at several scales, so when "bursts" (sudden high intensities of a given tag in space or time) are found, they become an indicator of event of meaningful place. In the Web-a-Where project, Amitay et al [8] associate web pages to geographical locations to which they are related, also identifying the main "geographical focus". The "tag enrichment" process thus consists of finding words (normally Named Entities) that show potential for geo-referencing, and then applying a disambiguation taxonomy (e.g. "MA" with "Massachusetts" or "Haifa" with "Haifa/Israel/Asia").

While our work focuses on the semantic aspect of location representation, we also take advantage of information available on the Web about public places. With the rapid growth of the World Wide Web, a continuously increasing number of commercial and non-commercial entities acquire presence on-line, whether through the deployment of proper web sites or by referral of related institutions. This presents an opportunity for identifying the information which describes how different people and communities relate to places, and by that enrich the representation of a Point Of Interest. Notwithstanding the effort of many, the Semantic Web is hardly becoming a reality, and, therefore, information is rarely structured or tagged with semantic meaning. Currently, it is widely accepted that the majority of on-line information contains unrestricted user-written text. Hence, we become dependent primarily on Information Extraction (IE) techniques for collecting and composing information on the Web.

Some relevant works can be referred in this realm, namely Open Calais [9] and Semantic Hacker [10], which focus on entity extraction from unstructured texts. They provide semantic indexes, although their focus is not restricted to information about space. A different approach, Scarlet [11], works on the extraction of the relations between concepts, an extremely challenging task within Information Extraction. Our approach shares references and methodologies with these works, but we emphasize information related to the place.

3 Automatic Labeling of a Point of Interest

In this section we present our system to semantically enrich POIs by Information Extraction techniques. This information can be obtained from different sources on the Web, what we call "perspectives" of a place. As we will see later these perspectives can be correlated and compared to induce different views according to what is intended to know about a place.

3.1 KUSCO

Having a set of pages as input, Kusco [1] extracts a ranked list of concepts. This process includes Noun Phrase chunking and Named Entity Recognition (NER) using available Natural Language Processing (NLP) tools [12,13]. *Noun Phrase chunking* is made typically by partial (sometimes called 'shallow') parsers and extract clusters of words that represent people or objects. They tend to concentrate on identifying *base* noun phrases, which consist of a *head* noun, i.e., the main noun in the phrase, and its *left modifiers*, i.e, determiners and adjectives occurring just to the left of it. Named Entity Recognition tries to identify proper names in documents and may also classify these proper names as to whether they designate people, places, companies, organizations, and the like. Figure 1 presents the overall architecture of our system.

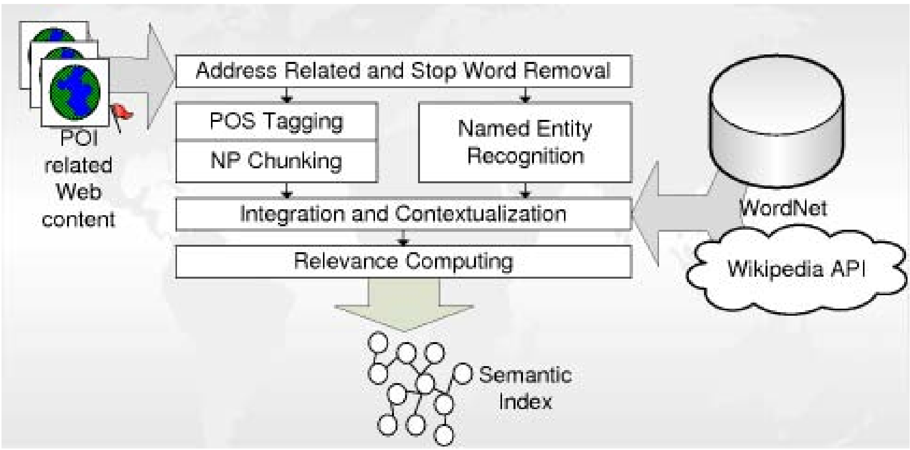


Fig. 1. Kusco architecture

On completion of these subtasks, for each page, KUSCO ranks the concept with TF-IDF [14] (Term Frequency × Inverse Document Frequency) value in order to extract the most relevant terms that will represent a given place. These nouns are contextualized on WordNet and thus can be thought not only as a word but more cognitively as a concept (specifically a synset - family of words having the same meaning, i.e., synonyms [15]). Given that each word present in WordNet may have different meanings associated, its most frequent sense is selected to contextualize a given term. For example, the term "wine" has two meanings in WordNet: "ermented juice (of grapes especially)" or "a red as dark as red wine"; being the first meaning the most frequent used considering statistics from WordNet annotated corpus (Semcor[16]).

When using data from different sources, integration of information is imperative to avoid duplicates. To solve this problem we treat differently common nouns (generally denoting concepts) from proper nouns (generally Named Entities found). Although we use WordNet to find synonyms in the first group, we don't have a list of all possible entities in the world to match words from the second group. So, we take advantage of the relatively mature field of *String metrics* to find the distance between strings using an open-source available library with different algorithms implementations [17].

3.2 The Perspectives

For any place, we build different lists of words, each list representing a "perspective" on the place, mostly dependent on the resource being analysed.

Open Web Perspective. The *Open Web* perspective consists of crawling the web using a search engine (Yahoo) given a POI name and address. The term "open" means that the search is not constrained to any particular web domain. The address is composed by the City name (where the POI is located) and is obtained from Gazetteers [1] available on Web. The search is made by the freely available YahooSearch API. We apply a heuristic that uses the geographical reference as another keyword in the search. Thus, assuming a POI is a tuple (Latitude, Longitude, Name), the final search query will be: <City Name> <Name>. To automatically select only pages centered on a given place, we filter out unuseful Web Pages with the following heuristics:(1) The title must contain the POI name; (2) The page body must contain an explicit reference to the POI geographical area; (3) Out of date pages will not be considered.

Wikipedia Perspective. Wikipedia provides us with a massive database of partially structured textual information, currently about over 3 million topics. Plenty of relevant information about places is obtainable, both directly by searching for the actual Wikipedia page of a POI (e.g. Starbucks), and indirectly by finding information related to its category (e.g. Restaurant). We now present the two variations currently implemented, the red Wiki (indirect approach) and the yellow Wiki (direct approach).

Low-precision labeling: the Red Wiki
In the Red Wiki perspective, we extract the Wikipedia page corresponding to the identified category of a POI. Local POI directories are normally structured in a hierarchical tree of categories. This taxonomy may be created by the company itself or be collaboratively built by suggestion of users who feed the system with new POIs. We don't assume a rigid organization neither a consistent validation of such taxonomy. So, node duplication and multiple heritance may be a reality that a generic methodology must face. In reality, in our database it is normal for each POI to have multiple categories.

Since no API is currently available from those local directories studied, we have created a wrapper based on regular expressions in order to automatically extract the category taxonomy of each local directory. Only Yelp web site provides the complete list of categories, while Yahoo Local only presents it through menu navigation along its web site. Curiously, this dynamics is also observed in the fact that this taxonomy is different depending on which city we are virtually visiting. Namely, Yahoo Local builds dynamically their menus, thus presenting proper taxonomies to distinct cities. Through time, this taxonomy grows with new types of services and places. In this way, by using specific wrappers to each

[1] A geographical dictionary generally including position and geographical names like Geonet Names Server and Geographic Names Information System [18].

POI provider, it is possible to run it periodically to integrate new categories in the respective stored taxonomy.

To contextualize each category in the corresponding Wikipedia article we base ourselves on string similarity between the category name and article title. We have opted for a top-down approach, from main categories to taxonomy leaves. To increase the confidence of this process, we disambiguate manually the main categories to start with and make sure that at least a more generic category will be connected to the Wikipages of its hypernym. When a POI has many categories, we obtain the articles for each one and consider the union of all the resulting articles as the source of analysis. Since there are many different combinations of categories, we can guarantee that each POI gets its own specific flavor of category analysis.

Medium-precision labeling: the Yellow Wiki
While the previous approach is centered on place category, here we focus our attention on Place name. We use string similarity to match Place name to Wikipage title in order to find the Wikipedia description for a given place. On a first glance, this method is efficient in mapping compound and rare place names such as 'Beth Israel Deaconess Medical Center' or 'Institute of Real State Management', however it can naively induce some wrong mappings for those places with very common names (e.g., Highway - a clothing accessories store in New York, Registry - a recruitment company in Boston, Energy Source - a batteries store in New York). We solved this problem by determining the specifcity of place names, and only considering those with high Information Content (IC)[19]. The Information Content of a concept is defined as the negative log likelihood, -logp(c), where p(c) is the probability of encountering such concept. For example, 'money' has less information content than 'nickel' as the probability of encountering the concept, p(Money), is larger than encountering the probability of p(Nickel) in a given corpus. For those names present in WordNet (e.g. Highway, Registry), IC is already calculated [16], while for those not present in WordNet, we heuristically assume that they are only considered by our approach if they are not a node in Wikipedia taxonomy, i.e., a Wikipage representing a Wikipedia category (case of Energy Source), but being only a Wikipedia article.

4 Experimental Results

We have collected a large set of Points of Interest from Boston, New York and San Francisco. The extraction of words for those POIs (to what we call *enrichment*) needs some processing time. The average for a POI analysis from the Open Web perspective is approximately 108 seconds, while Red and Yellow Wiki are 57 and 31 seconds, respectively. The Open Web is naturally more time consuming since it searches the entire web (using Yahoo search engine), while any of the other perspectives uses a more bounded search. In Table 1, we present the overall statistics.

Regarding the words obtained, we have a total of 77558 different words, of which 9746 (12.6%) were also identified in WordNet. An analysis to these concepts was made regarding the average information content (IC) obtained. The

Table 1. Above: number of POIs per perspective/city; Below: number of enriched POIs per perspective/city

	New York	**Boston**	**San Francisco**	**Overall**
Yahoo	183144	64133	94466	341743
YellowPages	7694	12878	-	20572
OpenWeb	757	2020	-	2777
Red Wiki	69011	20309	-	89320
Yellow Wiki	4400	1928	-	6328

IC [19] reflects the balance of specificity of the concept in a scale of 0 to 17. This average is 16.313395 (st.dev.=1.7263386), meaning that the concepts are in general very specific, thus carrying a rich content to the definition of POIs. This is however a risky game: if concepts are generic, the probability of being correct with respect to the place is much higher than when they are very specific. Since these words come from the actual text, in general they should be correct.

We show in Table 2 an excerpt of the "good" and "bad" examples found. This choice was made by the authors and intends to reveal the qualities and problems of the approach. A less subjective perspective on the results is presented in the next section. Except for the Red Wiki, we only put one category for each POI (many of them have more than one) to make the table more legible and let the reader understand the type of place.

The results from Open Web are extremely dependent on the initial search accuracy. In other words, if the correct webpage about the POI is found, then generally the results are acceptable, however this is not always simple to guarantee, depending on the nature of the POI. For example, if its name is a common noun (e.g. "Gap"), there will be too many unrelated pages, if it doesn't have a webpage (e.g. it only exists in directory listings), there won't be any page. The Red Wiki perspective easily obtains meaningful words, although hardly specific to the POI, which is expectable since it works on its category. It is therefore a very "safe" perspective in terms of guaranteeing correctness of the obtained semantics. The Yellow Wiki can get much more refined results (e.g. the TD Garden is in fact the arena where Boston Celtics play NBA games) but it is more fragile when the wrong Wikipedia page is found (e.g. the Blue Smoke stake house is taken as a film with the same name) and is easily fooled by lateral information (e.g. Starbucks being from Seattle should be less relevant than for example for serving coffee or cappuccino).

5 Validation

We face an important challenge of understanding the actual quality of the results in terms of the *correctness* of the words assigned to places. The *ideal* list of words is by nature subjective. As referred above, a place can be defined according to different perspectives, and each perspective can vary with subject. In terms of validation, this raises difficult questions even for the typical user

Table 2. Some examples from experiments (in each perspective, first 2 are "good" examples, last 2 "bad" examples)

Name	**Categories**	**Terms**
Open Web		
Envirotech Incorporated	Waste and Environmental Consulting	Industrial Services, Asbestos Management, Mildew Removal, Asbestos Removal, Residential Services
Grasshopper	Telecommunications	Boston Telecommunications, Gary, Communication Services, Boston Business Directory, Telephone Communications
I Have A Dream Foundation	Educational Consulting	Boulder County, Dany Garcia, Arne Duncan, Jeffrey Gural, National Partners
Monroe Paint Distributors Incorporated	B2B Paint & Wall Coverings	movie theater, latitude, beauty salon, Delicious, Construction
Red Wiki		
Harvard Market	Grocery Stores	groceries, retailing, food, vegetables, products
Kim Depole Design Incorporated	Interior Design	office space, architects, private residence, code, decoration
Cambridge Library	Libraries	collection, library, information needs, public body, access points
Harvard Magazine	Marketing Agencies, News Services	pool, product, industry trade group, farmers, consumers
Yellow Wiki		
Boston Police Department	Law Enforcement	Massachusetts, law enforcement agency, correction, investigation, responsibility
TD Garden	Entertainment Venues	Boston Celtics, arena, Boston Blazers, naming rights, National Lacrosse League,
Starbucks Coffee	Coffee Houses	stores, Seattle, Washington, drip, Israeli
Blue Smoke	Steak Houses	Nora Roberts, Blue Smoke, Television film, novel

survey. A very large sample of people that *know* the specific places is necessary to achieve believable results, which then becomes unpractical and costly. We decided to analyze our results according to 2 dimensions: category consistency and coherence among perspectives. We also analyze the distribution of the words in each perspective.

Each POI has ultimately one category[2], so, in the *category consistency* validation, the task is to verify the stability of the word patterns according to those categories (15 for POIs). The first approach is to apply a clustering algorithm such as K Means, where K corresponds to the number of different categories. After clustering with a training set, we apply a classification task: given the 5 top words of a POI from the test set, classify the POI in one of the categories. We apply 10 fold cross validation[3]. In order to get basic benchmarks to analyze the results, we set up two baselines: the *random baseline* consists of the accuracy of a random classifier (applied to all cases of the data set); the *fixed baseline* classifier selects the most popular class.

The resulting accuracy of clustering is in fact extremely poor, even when compared with the baselines. The highest value obtained (37.66%) was for the OpenWeb perspective which is actually lower than the *fixed baseline* of 47.49%(the accuracy obtained by a dumb model which basically always assigns the most popular category to any POI). This implies that either the word patterns are not constant with respect to category or they are more elaborate than achievable with clustering algorithms. We tried Bayesian Networks, which are actually more common in text categorization, and the results improved considerably (accuracy from 57.12% to 97.3%). In Table 3, we summarize the results. The high value for the Red Wikipedia perspective reflects that our algorithm could extract sufficiently specific words from Wikipedia category definitions such that they become easily distinguishable form each other. This is interesting because many POIs (those that had multiple categories) were assigned a more generic category for classification, thus gathering in the same *class* POIs from different original areas of the category hierarchy and with many different words. Taking this into account, we can conclude that the original assignment of categories to the POIs is itself very consistent (e.g. Food & Dining subcategories are rarely mixed with Home & Garden ones).

For the Yellow Wikipedia perspective, the patterns are still extremely stable while, for Open Web , the results become less prominent. For Open Web a careful analysis reveals that there is still a reasonable quantity of noise in the indexes as we are collecting information from different Web sites with distinct templates and lateral information (e.g. advertisments, ads by Google, news headlines from RSS feeds) while in Wikipedia we have an available API to extract only useful

[2] In reality, only a portion of the POIs have a single category, but we determined the *Least Common Subsummer* for POIs with multiple categories in the hierarchy, which consists of the most specific upper category that contains the categories of the POI in its descendants.

[3] Divide data set into 10 folds, each fold will become a test set to a model built with the remaining 9 folds[20].

and structured information. We can thus conclude that, for all perspectives, our system brings a degree of consistency that is relevant, particularly considering the two baselines.

Table 3. Category consistency results

	Rand. baseline	**Fixed baseline**	**K Means**	**Bayesian Network**
Red Wiki	9.199%	16.54%	24.40%	97.3%
Yellow Wiki	13.483%	23.25%	28.26%	66.56%
Open Web	23.781%	42.49%	18.71%	57.12%

The Best Case of similarity is the closest pair to a given POI, and the Worst Case the farthest one. The overall Best and Worst cases are detailed in Table 4. Terms mapped into WordNet are complemented with synonyms enclosed by parentheses, "()". Looking at the examples the dispersion between perspectives may be not a disadvantage but a richness acquired by the contribution of distinct terms from different perspectives. And maybe even for concepts not exactly synonymous but related (e.g, telecommunications, telephone) we can apply in the future other semantic similarity measures taking in account the meaning of each particular concept.

Table 4. An example of High similarity (above) and another of Low similarity (below)

Dexter School	(categories: Parochial Schools, Elementary Schools, Middle Schools, High Schools, Preschools)
Perspective	**Terms**
Yellow Wiki	Grade(class, form, course), students, Schools, teamwork, Francis Caswell,...
Red Wiki	compulsory education, tuition(tuition fee), teachers, North America, students,...
Grasshopper	(category: Telecommunications)
Open Web	Boston Telecommunications, Gary, Communication Services, Boston Business Directory, Telephone Communications,...
Red Wiki	Telecomunication(telecom), modern times, telephone(phone, telephone set), inventors, semaphore,...

We also randomly produced a sample of 420 Semantic Indexes (Red and Yellow Wiki) about Boston POIs which were manually validated by 28 volunteers who know the city in study, answering the question, for each word, whether it is related to the POI or not. We obtained a precision of 58% ($\sigma = 15\%$) and 56% ($\sigma = 20\%$) for Yellow and Red Wiki perspectives, respectively, considering all unanswered tags as invalid. In some cases even the volunteers disagree, reflecting the subjective nature of this information. Finally, we also check the shape of the word frequency histogram. In every perspective, we observed the distribution of words follows the typical long tail distribution that matches Zipf's law for word frequency [21], which was an expectable result.

6 Conclusions and Further Work

In this paper, we presented our work on semantic annotation of places from web resources. The implemented system could gather a massive amount of POIs and analyze a large portion of it, clearly enough for a valid analysis. The experiments show that the semantic indexes obtained have an average good quality, and we presented several different "perspectives" that can be used according to the context. When we need to guarantee the correctness of the words, we should use the Red Wiki perspective, which sacrifices the specificity of each POI for the analysis of its category (which brings normally correct results). When we are looking for exact information about a specific place, we can use the Yellow Wiki or the Open Web perspective. The former is preferable when the POI exists in the Wikipedia while the latter is the only option otherwise.

To achieve a refinement of our system, we plan to implement a High-Precision place labeling for Wikipedia articles increasing the realiabity of Yellow Wiki, namely for specific names (with high Information Content) that are not related at all with the Place we analyze. For example, 'Apostrophe' is a store for commercial photographers in San Francisco, however searching in Wikipedia, we obtain the page describing the punctuation mark. This seems a gap in our methodology that could be solved by comparing the similarity between the place index with their corresponding category (Commercial Photographers). Another confidence level we can add to this mapping could be to verify if the geographic information contained in Wikipedia pages are related to a given place, be it by geo-reference annotations (presently available in more than 500000 articles[4]) or by textual patterns (country, city or neighborhood name).

As an application of our work, an interesting further step consists of trying to deduce semantic correlations between places and create a model of land use to understand the pattern of business/human occupation in some regions in the city. While others interested researchers could use our semantic enriched data about POIs over a given city using directly available API web services to query through different perspectives[5].

References

1. Alves, A., Pereira, F.C., Biderman, A., Ratti, C.: Place enrichment by mining the web. In: Proc. of the Third European Conference on Ambient Intelligence (2009)
2. Genereux, R., Ward, L., Russell, J.: The behavioral component in the meaning of places. Journal of Environmental Psychology 3, 43–55 (1983)
3. Kramer, B.: Classification of generic places: Explorations with implications for evaluation. Journal of Environmental Psychology 15, 3–22 (1995)
4. Harrison, S., Dourish, P.: Re-place-ing space: the roles of place and space in collaborative systems. In: CSCW 1996, pp. 67–76. ACM Press, New York (1996)

[4] By Wikipedia Geoname Web Server http://www.geonames.org/export/Wikipedia-webservice.html

[5] Available for visualization at
`http://greenhomes.dei.uc.pt:8080/Maps/index.html`

5. Hightower, J.: From position to place. In: Proc. of the Workshop on Location-Aware Computing (2003); 10–12 part of the Ubiquitous Computing Conference
6. Lemmens, R., Deng, D.: Web 2.0 and semantic web: Clarifying the meaning of spatial features. In: Semantic Web meets Geopatial Applications, AGILE 2008 (2008)
7. Rattenbury, T., Good, N., Naaman, M.: Towards automatic extraction of event and place semantics from flickr tags. In: SIGIR, USA, pp. 103–110. ACM, New York (2007)
8. Amitay, E., Har'El, N., Sivan, R., Soffer, A.: Web-a-where: geotagging web content. In: SIGIR 2004, pp. 273–280. ACM, New York (2004)
9. Reuters, T.: Open calais api (2009)
10. TextWise: Semantic hacker api (2009)
11. Sabou, M., d'Aquin, M., Motta, E.: Exploring the semantic web as background knowledge for ontology matching. Journal of Data Semantics (2008)
12. Ramshaw, L., Marcus, M.: Text Chunking using Transformation-Based Learning. In: WVLC, Cambridge, USA (1995)
13. Finkel, J.R., Grenager, T., Manning, C.: Incorporating non-local information into information extraction systems by gibbs sampling. In: ACL 2005, pp. 363–370 (2005)
14. Salton, G., Buckley, C.: Term-weighting approaches in automatic text retrieval. Information Processing and Management 24(5), 513–523 (1988)
15. Fellbaum, C.: WordNet: An Electronic Lexical Database (Language, Speech, and Communication). The MIT Press, Cambridge (1998)
16. Mihalcea, R.: Semcor semantically tagged corpus. Technical report, CiteSeer (1998)
17. Cohen, W.W., Ravikumar, P., Fienberg, S.E.: A comparison of string distance metrics for name-matching tasks. In: IJCAI WS Inf. Integration, pp. 73–78 (2003)
18. Imagery, N., Agency, M.: Geonet names server (gns) (2007)
19. Resnik, P.: Using information content to evaluate semantic similarity in a taxonomy. In: IJCAI, pp. 448–453 (1995)
20. Witten, I.H., Frank, E.: Data Mining: Practical machine learning tools and techniques, 2nd edn. Morgan Kaufmann, San Francisco (2005)
21. Zipf, G.K.: Selective Studies and the Principle of Relative Frequency in Language. Harvard University Press, Cambridge (1932)

AmbiSec: Securing Smart Spaces Using Entropy Harvesting

Paolo Barsocchi[1], Stefano Chessa[1,2], Ivan Martinovic[3], and Gabriele Oligeri[1]

[1] ISTI-CNR, Pisa Research Area, Pisa, Italy
[2] Computer Science Department, University of Pisa,
Largo B. Pontecorvo 3, 56127 Pisa, 8 Italy
[3] Distributed Computer Systems Lab, Computer Science Department,
University of Kaiserslautern, Germany

Abstract. Following the vision of Ambient Intelligence (AmI), this paper introduces and evaluates a novel security scheme that takes the advantage of the unpredictable and erratic behavior of wireless communication to generate secret keys. The main advantage is that the secret key generation is applicable to every wireless device, independently of their hardware characteristics as it only requires a wireless interface and a human movement, which inherently affects the signal propagation within the physical environment. To analyze the applicability of this scheme, we implement and systematically evaluate the key generation using a wireless sensor network deployed in a real-world scenario. The analysis clarifies how different factors influence the amount of randomness collected from the physical environment, and it also shows that guessing attacks from an eavesdropper are negligible even if it is able to eavesdrop the complete wireless communication.

1 Introduction

Ambient Intelligence (AmI) refers to digital environments that proactively support people in their daily activities. In the AmI context, the European Commission recently started the Ambient Assisted Living (AAL) technology and innovation funding program. This program aims at improving the quality of life of the elderly persons, by increasing their autonomy, assisting them in their daily activities, and by enabling them to feel included, secure, protected and supported. AAL spaces are physical places featured with AmI enabling technologies, including intelligence that supports various services. Examples of AAL spaces are, e.g., a home where a user lives, or a place where the user works, and even public buildings such as hospitals or airports, or even a whole town. At a smaller scale an AAL space may also comprise the body of the user itself. For the creation of AAL spaces, it is essential to have a framework supporting context-awareness. The scope of the contextual information spans the situational user context (his or her identity, capabilities, preferences, etc.) and environmental parameters that are usually collected by a large number of wireless sensors spread in the environment and on the user's body. Such sensors typically form a

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 73–85, 2010.

Wireless Sensor Network (WSN) [3], in which some of the sensors are statically deployed in the environment, while others are mobile (typically those deployed on the user's body). The sensors are also connected to a gateway interconnecting them with the application framework of the AAL system [5]. Under this aspect the static, environmental sensors provide a communication backbone that connects mobile sensors to a gateway.

One of the major issues in this context is the security of communication within the WSN. In fact, the sensors collect a large amount of confidential data related to the activities of the user and to his/her physiological or emotional state.

Furthermore, an attacker tampering with such data may easily impair the AAL space functions and may, possibly, even cause a greater damage (for example, an attack could result in wrong or missed medications/therapies). On the other hand, sensors communicate wirelessly, hence communications can be easily eavesdropped or tampered with. Furthermore, since wireless sensors are commonly highly computational and battery-constrained, conventional solutions based on asymmetric cryptography are usually not feasible. Similarly, the use of pre-assigned, symmetric keys may be impaired by attackers that tamper with the sensors' hardware. To circumvent these problems, the sensors are typically pre-assigned with a symmetric, secret key that is periodically refreshed. The transmission of a fresh key to the sensors clearly needs to be encrypted, however, if a sensor has been compromised, the fresh key will also be compromised. In the effort of making the secret key generation and distribution more secure, a recent approach that achieved promising results exploits the observation of radio wave propagation and its erratic and unpredictable behaviors in temporal and spatial domain (such a process of observing the environmental behavior to the purpose of a secret key generation, we refer to as *entropy harvesting*). With this approach, a pair of sensors may establish a new, secret key based on the observation of the radio waves propagation and by leveraging the reciprocal nature of the wireless channel. The matching of the keys computed independently by the two sensors can be ensured since the behavior of the wave propagation is highly correlated. On the other hand, an attacker observing the communication between the two sensors will observe a different channel behavior and will thus be unable to determine the same key generated by the sensors. Although this approach can be implemented by any pair of sensors without particular requirements and difficulties, it has two main drawbacks: (i) the channel behavior may not be variable enough to provide the entropy necessary for the secret key generation, and (ii) if the attacker is too close to one of the sensors it may observe the same channel behavior making it possible to guess the key. These two issues have been the subject of recent and intensive studies summarized in Section 2.

In this paper, we reconsider the secret key generation based on entropy harvesting in AAL scenarios. In particular, we focus on the generation of secret keys among mobile and static sensors. The static sensors are a part of the infrastructure of an AAL system, and they typically have less constraints than mobile sensors which are commonly applied on the body of a person. Hence, the static sensors may implement more complex protocols for secret key distribution/generation and they may even be subject to constant monitoring and maintenance (as it is expected if such sensors constitute an AAL facility in a public environment as, for example, an office or a hospital). On the other hand, mobile sensors may dynamically enter/exit the AAL space since they could be deployed on an elder or impaired person who is typically unable to manage them. For these reasons the mobile sensors are more vulnerable to attacks, yet their security becomes crucial. Assuming static sensors as safe, the process of entropy harvesting greatly benefits (in terms of key generation speed) from the randomness induced by the user's body mobility. Moreover, we show that this scenario is even more challenging for an attacker observing a wireless channel in proximity of a mobile or a static sensor, since the body of the user and his/her mobility alter the wireless channel in an unpredictable way even at short distances.

2 Related Work

Several papers take advantage of the wireless channel unpredictable behavior and use narrow-band communication to generate secret keys from it. In [7], the authors assume random movement of transmitters as the source of entropy. By frequently sampling the received signal strength, both parties can create a sequence of channel states which are strongly correlated. Similarly, in [2], the authors propose a protocol that focuses mainly on the robustness of the key generation process, i.e., tolerance against deviations of the wireless channel and a high success rate of key-agreements. Their quantitative evaluation is based on a theoretical channel model and simulations which, however, do not shed much light on the performance of such key-generation schemes in an everyday, real-world environment. Several other contributions use a highly specialized hardware, such as steerable antennas, ultra-wideband (UWB) radio or multi-antenna systems with performance-capable processors [1,10,11]. The most recent overview with comparison of existing schemes is given in [6]. Although these existing results justify the possibility of generating secret keys from wireless communication, they still cannot be applied in static WSNs. Commodity sensor devices can neither rely on random (and fast) movements, nor on sophisticated hardware, such as, steerable multi-array antennas. The authors in [9] present a similar scenario where they also consider static WSNs. In fact, their key-generation scheme is based on using frequency-hopping as a source of randomness. However, a limited number of very narrow wireless channels is not suitable to generate long secret keys.

3 Scenario

The envisioned scenario is constituted by three main entities: the user $\mathcal{U}$, the adversary $\mathcal{E}$, and finally the smart space. The user $\mathcal{U}$ requests a secure communication with the smart space to avoid $\mathcal{E}$ being able to violate her privacy.

The smart space: It is basically constituted by an intelligent environment providing services to the user $\mathcal{U}$. The smart space is connected to Internet, and characterized by Secure Network Infrastructure (SNI) which, in turn, is constituted by B sensors $\{\mathcal{B}^0, \mathcal{B}^1, \ldots, \mathcal{B}^{B-1}\}$ (also called anchors). We assume that all the anchors can communicate among each other in a secure way as part of SNI, i.e., communication between $\mathcal{B}^i$ and $\mathcal{B}^j$ is secure for each $0 \leq i, j < B$.

The user: She is provided with the sensor $\mathcal{A}$, which in turn communicates with SNI. Our solution relies on the *on-board RSSI estimation capability* of $\mathcal{A}$, i.e., $\mathcal{A}$'s radio provides an estimation of the received signal strength for each received packet.

We assume that $\mathcal{A}$ is pre-loaded with a secret key $\mathbf{K}_0$, which allows her to be authenticated by SNI, and start to communicate with SNI in a safe fashion.

The adversary: The adversary $\mathcal{E}$ is interested to know the entire secret information of $\mathcal{U}$. Our envisioned adversary is a global eavesdropper on $\mathcal{A}$ and commits on the *honest but curious* behavior, that is, even if she is able to physically break-in $\mathcal{A}$, she never injects code on it or, more generally, changes the original behavior of $\mathcal{A}$. $\mathcal{E}$'s behavior is justified by the fact that she wants to stay hidden to collect as many information on the user as possible. Therefore, code injection strategies should be avoided from the point of view of the adversary, since an active adversary who changes the sensor's behavior may eventually be discovered by SNI.

We consider two possible $\mathcal{E}$'s behavior type: (i) $\mathcal{E}$ places at least one stealthy sensor on $\mathcal{U}$, then she behaves like a global eavesdropper and consequently, she gathers all the transmitted data from $\mathcal{A}$ and attacks it by an exhaustive brute-force; (ii) $\mathcal{E}$ breaks-in $\mathcal{A}$, that is, the overall memory content becomes clear to $\mathcal{E}$. It's worth noticing the differences between the two scenarios: while in the first case $\mathcal{E}$ could eventually guess the shared secret between $\mathcal{A}$ and SNI, in the second case, following the compromise phase, $\mathcal{E}$ knows all the $\mathcal{A}$'s secrets, and being a global eavesdropper on the environment, she is able to mimic all the crypto functions running on $\mathcal{A}$. Therefore, in this scenario, crypto algorithms are useless against the adversary, but the source of a "fresh" randomness, i.e., data unknown to $\mathcal{E}$, turns out to be a good solution for regaining security after the compromise event, even if $\mathcal{E}$ is still a global eavesdropper on $\mathcal{A}$.

3.1 Definitions and a Baseline Scenario

The baseline scenario is constituted by one only anchor, hereinafter $\mathcal{B}^0$, and a mobile user $\mathcal{U}$ provided with the sensor $\mathcal{A}$. In this scenario we make the following assumptions:

RSSI estimation capability: The core of our system is the RSSI evaluation procedure. Sensors must be able to evaluate the received signal strength (hereinafter RSSI) for each received packet.

Security: Each sensor is able to perform symmetric key encryption and hash computations. There is a pre-shared $\mathbf{K}_0$ secret which is injected in $\mathcal{A}$ before the sensor is placed on the user $\mathcal{U}$.

$\mathcal{A}$ **behavior:** It performs one environmental sensing per round, encrypts the sensed data according to our proposed algorithm, and finally transmits it to the SNI.

$\mathcal{U}$ **behavior:** The user is walking round in an office environment, with the $\mathcal{A}$ and $\mathcal{E}$ sensors on her body.

$\mathcal{E}$ **behavior:** The adversary is a global eavesdropper on the environment, and collects all the received and transmitted packets by $\mathcal{A}$.

$\mathcal{A}$ senses the user information in each round t, and then transmits the cyphered data to $\mathcal{B}^0$, which in turn, acknowledge it. Let us define r^0_t as the RSSI estimated by $\mathcal{B}^0$ for the packet sent by $\mathcal{A}$ to $\mathcal{B}^0$ in the round t. Let us also define r^a_t as the RSSI estimated by $\mathcal{A}$ for the packet sent by $\mathcal{B}^0$ to $\mathcal{A}$ in the round t. Our secret key generation relies on the assumption that $r^a_t \approx r^0_t$, i.e. $\mathcal{A}$ and $\mathcal{B}^0$ have a shared secret constituted by the received power evaluated on two subsequent packets that they have exchanged with oneanother. The main idea of our approach involves: (i) to correct small deviations between r^a_t and r^0_t without exchanging information between $\mathcal{A}$ and SNI, (ii) to collect a sufficient number of shared secrets in order to make a secret key, and finally, (iii) to commit on the shared secret.

The proposed key generation and agreement protocol is composed of three phases, namely the *sampling phase*, the *symbol generation phase*, and the *key agreement phase*. Figures 1 and 2 show the outline of the sampling and symbols generation phases at both $\mathcal{A}$ and $\mathcal{B}^0$ sides.

$$\mathcal{A}: \underbrace{\underbrace{r^a_0,\dots,r^a_{w-1}}_{S_0}\underbrace{r^a_w,\dots,r^a_{2\cdot w-1}}_{S_1}\cdots\underbrace{r^a_{(L-2)\cdot w},\dots,r^a_{(L-1)\cdot(w-1)}}_{S_{L-1}}}_{\mathcal{K}_1}$$

Fig. 1. Sampling RSSI, symbols and key generation at $\mathcal{A}$ side

$$\mathcal{B}^0: \underbrace{\underbrace{r^0_0,\dots,r^0_{w-1}}_{S_0}\underbrace{r^0_w,\dots,r^0_{2\cdot w-1}}_{S_1}\cdots\underbrace{r^0_{(L-2)\cdot w},\dots,r^0_{(L-1)\cdot(w-1)}}_{S_{L-1}}}_{\mathcal{K}_1}$$

Fig. 2. Sampling RSSI, symbols and key generation at $\mathcal{B}^0$ side

During the **sampling phase**, both $\mathcal{A}$ and $\mathcal{B}^0$ collect w RSSI samples, i.e., $\{r^a_0,\dots,r^a_{w-1}\}$ and $\{r^0_0,\dots,r^0_{w-1}\}$, respectively. It's worth noticing that these values may not be exactly the same, i.e., it may happen that $r^a_t \neq r^0_t$, for each $t \geq 0$. For this reason the symbol generation phase implements a correction algorithm which allows to obtain from the sequences $\{r^a_0,\dots,r^a_{w-1}\}$ and $\{r^0_0,\dots,r^0_{w-1}\}$ the same symbol S^1_0.

The **symbol generation phase** first computes the mean on the last w collected RSSI values at both sides, e.g., during the first round $\mathcal{A}$ and $\mathcal{B}^0$ compute $e_0^a = E\{r_0^a, \ldots, r_{w-1}^a\}$ and $e_0^0 = E\{r_0^0, \ldots, r_{w-1}^0\}$, respectively. Second, the mean values e_0^a and e_0^0 are quantized on a pre-constituted set of bins of width b_w for each one, see Fig. 3. The latter procedure allows to commit with high probability on the same symbol sequence by filtering the fast fluctuations due to the multipath fading effect, see for example rounds from which symbols S_2 and S_3 are generated in Figure 3.

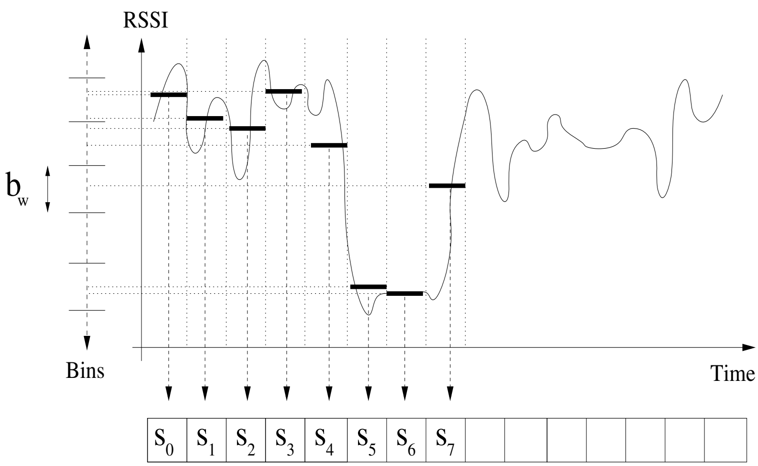


Fig. 3. Symbol generation phase at $\mathcal{A}$ and $\mathcal{B}^0$

The **key generation phase** collects the last L generated symbols and build up the key, e.g. at the end of the first round $\mathcal{A}$ and $\mathcal{B}^0$ have the keys $\mathcal{K}_0^0$ and $\mathcal{K}_0^a$, respectively.

The **key agreement phase** is used to check for the actual agreement on the computed keys at both parties, i.e. $\mathcal{K}_0^0 = \mathcal{K}_0^a$. The agreement test relies on the comparison of the hashes computed on the previously generated keys at both parties, i.e. both $\mathcal{A}$ and $\mathcal{B}^0$ compute $\mathcal{H}^a = H(\mathcal{K}^a)$ and $\mathcal{H}^0 = H(\mathcal{K}^0)$ on their generated keys $\mathcal{K}^a$ and $\mathcal{K}^0$, respectively; then, they transmit to each other $\mathcal{H}^a$ and $\mathcal{H}^0$, and verify the agreement.

3.2 Multi-anchors Scenario

In a realistic smart space, SNI usually provides more than one anchor B^0. For this reason we consider B different anchors: $\{\mathcal{B}^0, \ldots, \mathcal{B}^{B-1}\}$. The sampling procedure, symbol and key generation phases are the same of Section 3.1, but the key generation phase is improved. According to the baseline scenario, every w rounds a new symbol is generated, while every Lw rounds $\mathcal{A}$ and $\mathcal{B}^0$ commit on a new shared key; in a multi-anchors environment $\mathcal{A}$ commits on B keys every Lw

rounds. The key sequence $\{\mathcal{K}_0^0, \ldots, \mathcal{K}_0^{B-1}\}$ is hashed to obtain the session key $\mathbf{K}_1$. In this way, the session key $\mathbf{K}_1$ is a one way function of all the keys established with all the anchors in the SNI. Hence, in this case, a non-trusted third party must guess all the keys established between $\mathcal{A}$ and each anchor to compute the session key.

Table 1 shows how the session key is generated in a 6-anchors environment. After Lw rounds, $\mathcal{A}$ establishes one shared key with each anchor in the SNI, i.e., $\{\mathcal{K}_0^0, \ldots, \mathcal{K}_0^5\}$. Hence, the session key $\mathbf{K}_1$ is obtained hashing all the previous established keys.

Table 1. Key generation in a multi-anchors environment

Anchors \ Rounds	$0 \ldots Lw-1$	$Lw \ldots 2Lw-1$	$2Lw \ldots 3Lw-1$	$3Lw \ldots 4Lw-1$	$\rightarrow$
$\mathcal{B}^0$	$\mathcal{K}_0^0$	$\mathcal{K}_1^0$	$\mathcal{K}_2^0$	$\mathcal{K}_3^0$	...
$\mathcal{B}^1$	$\mathcal{K}_0^1$	$\mathcal{K}_1^1$	$\mathcal{K}_2^1$	$\mathcal{K}_3^1$	...
$\mathcal{B}^2$	$\mathcal{K}_0^2$	$\mathcal{K}_1^2$	$\mathcal{K}_2^2$	$\mathcal{K}_3^2$	...
$\mathcal{B}^3$	$\mathcal{K}_0^3$	$\mathcal{K}_1^3$	$\mathcal{K}_2^3$	$\mathcal{K}_3^3$	...
$\mathcal{B}^4$	$\mathcal{K}_0^4$	$\mathcal{K}_1^4$	$\mathcal{K}_2^4$	$\mathcal{K}_3^4$	...
$\mathcal{B}^5$	$\mathcal{K}_0^5$	$\mathcal{K}_1^5$	$\mathcal{K}_2^5$	$\mathcal{K}_3^5$	...
Hashing	$\downarrow H(\circ) \downarrow$	$\downarrow H(\circ) \downarrow$	$\downarrow H(\circ) \downarrow$	$\downarrow H(\circ) \downarrow$	...
Session Keys	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	...

4 Securing Smart Spaces

The adversary can guess the session key $\mathbf{K}_t$ in four different ways:

Brute-forcing the on-fly packet: The adversary accesses the session key $\mathbf{K}_t$ and the data carried by the current packet. Nevertheless, data sent before the compromise event is still safe because the session key $\mathbf{K}_{t-1}$ cannot be guessed from the broken key $\mathbf{K}_t$, subsequent packets are safe as well: the session key $\mathbf{K}_{t+1}$ cannot be computed from the current broken one.

Break-in the sensor: This is a more powerful attack with respect to the previous one; in this case the adversary breaks-in the sensor at round t and discovers all the secrets of the sensor for the current round t, but again, our protocol is able to strongly bound the adversary action. As in the previous case, in fact, after the adversary releases the sensor at round $t+1$, the sensor immediately regains its secret state, although the adversary knows everything about the sensor and it is a global eavesdropper on the environment.

RSSI guessing attack: This attack aims at guessing the RSSI values in order to reconstruct the shared secret between $\mathcal{A}$ and the anchors. Actually, a closer position of $\mathcal{E}$ with respect to $\mathcal{A}$ can help the adversary to guess the key, i.e., the radio of $\mathcal{E}$ experiences the same RSSI behavior of $\mathcal{A}$. To perform this attack, the adversary must guess all the pair-wise keys between $\mathcal{A}$ and the anchors.

Statistical analysis of the symbols: Session keys could be also guessed performing a statistical analysis of the RSSI values. Only uniform distribution of the RSSI values can give us real safe keys, i.e. the RSSI value at round t does not give any information about the value of the RSSI at round $t+1$ or $t-1$ (channel with no memory). The previous is a strong assumption, in fact the received signal strength has time correlation, but people movement can help us to reduce the channel coherence. In Section 5.3 we show a statistical description of our generated symbols and we give an estimation of how far the distribution of the generated symbols is from the "perfect" one (i.e., the uniform distribution), and finally, we provide a measure of the entropy related to the generated keys.

4.1 Dealing with Packet Loss

Packet loss must be taken into account in a real indoor environment, in particular in Smart Spaces where a lot of disturbance effects are present, i.e., interferences due to other WLANs, people movement, electromagnetic fields, etc. The baseline scenario constituted by the $\mathcal{A}$ sensor and just one anchor $\mathcal{B}^0$, could be significantly affected by the packet loss phenomena, i.e., if only one packet is lost, its RSSI estimation cannot be performed, and consequently, symbol and key generation phases fail.

In a real multi-anchor scenario more than one anchor is provided. In this case, no matter if one or more keys cannot be used to generate the session key due to failed agreements. The sensor $\mathcal{A}$ and the set of anchors belonging to SNI $(\mathcal{B}^0, \ldots, \mathcal{B}^{B-1})$ know the key agreements, and therefore, they can use the right keys to generate the session key. For example, recalling Table 1, let us assume $\mathcal{K}_0^1$ and $\mathcal{K}_0^4$ failed the agreement phase due to one or more lost packets, the session key $\mathbf{K}_1$ will be a one way function of the other ones, i.e., $\{\mathcal{K}_0^0, \mathcal{K}_0^2, \mathcal{K}_0^3, \mathcal{K}_0^5\}$. It is worth noticing that by using this session key the construction strongly improves the security robustness of our system, that is, even if the adversary eventually is able to guess all the keys for the current round, it also has to experience exactly the same packet loss of the $\mathcal{A}$ sensor.

5 Experimental Results

In our measurements we used the MICAz sensor motes [4], which are equipped with a 2.4 Ghz radio, IEEE 802.15.4 compliant. MICAz motes feature the Atmel ATmega128L low-power microcontroller and ChipCon radio CC2420. We performed several experimental measurements in order to evaluate the system performance in terms of key agreement and guessing key frequency. We conducted our experiments in the offices of the ISTI-CNR in Pisa, during a common work

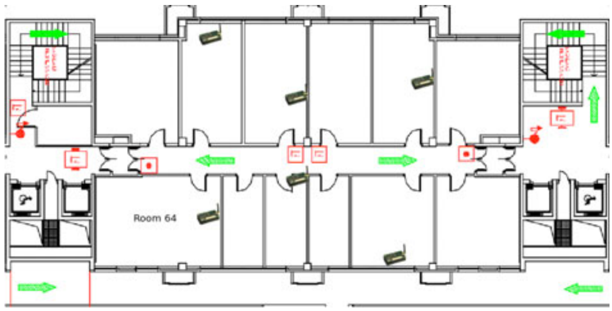


Fig. 4. The offices plan with the SNI anchors location

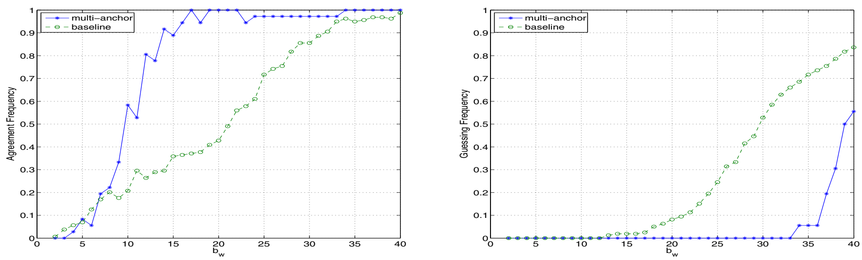
day. The $\mathcal{A}$ and $\mathcal{E}$ sensors were put on the neck and on the wrist, respectively, of the user $\mathcal{U}$. The SNI is constituted by 6 anchors deployed as shown in Figure 4. The environment consists of adjacent small office rooms and a corridor located on the second floor of ferroconcrete building whose windows face a neighboring building with metal exterior walls.

We consider two different measurement scenarios: the baseline where the user $\mathcal{U}$ is moving around in an office environment (Room 32 in Fig. 4), while $\mathcal{A}$ sensor is communicating with just one anchor $\mathcal{B}^0$; and a real multi-anchor scenario constituted by 6 sensors, where the user and other people are moving around the entire map. In order to study the effectiveness and robustness of our approach we introduce two new metrics, the *key agreement frequency*, hereafter F_a, and the *key guessing frequency*, hereinafter F_g. While the former is used to measure the effectiveness on the generation of the session key between $\mathcal{A}$ and SNI, the latter estimates the adversary key guessing capabilities, i.e. how frequently the adversary is able to see the same RSSI value of $\mathcal{A}$, and consequently, guessing the session key.

5.1 Key Agreement and Guessing Performance

In this subsection we will show the performance results in terms of key agreement frequency F_a and key guessing frequency F_g. In all the simulations we used $w = 5$, i.e. during the symbol generation phase we computed the mean over the last 5 RSSI values. Figure 5 shows the performance results as a function of the bin width b_w for the baseline and the multi-anchor scenarios. Key agreement is more frequent with larger bins; in particular, when $b_w = 20$, the key agreement frequency reaches 0.4 for the baseline scenario, while it reaches 1 for the multi-anchor scenario. Guessing key performance has the same behavior as function of the bin width.

It is worth noticing that choosing $b_w = 20$, yields $F_g = 0.1$ for baseline scenario, while $F_g = 0$ for the multi-anchor scenario. Hence, choosing a $16 < b_w < 33$ our algorithm guarantees a key agreement frequency between $\mathcal{A}$ and the SNI, avoiding the RSSI guessing attack in a multi-anchors scenario.



(a) Key agreement frequency. (b) Key guessing frequency.

Fig. 5. Performance in a baseline and multi-anchor scenario

5.2 Key Length Analysis

Key length is an important issue to address; longer keys need more symbols to be generated, and as consequence, more time. On the contrary, shorter keys are fast to obtain but could be unsafe, i.e., they may have insufficient randomness. Figure 6 shows the key agreement frequency and the key guessing frequency as function of the key length. We used key lengths in the range from 8 to 128 bits, and we analyze the algorithm effectiveness in terms of key agreement (F_a) against the robustness to the adversary RSSI guessing attack (F_g). The longer keys are more difficult to generate, since the channel is supposed to be symmetric for longer time intervals. The key agreement frequency is almost 1 for the shortest key (8 bit) and decreases to ≈ 0.9 for the longest one (128 bit). The key guessing frequency has the same trend but it is steeper respect to the key agreement frequency, i.e. a longer key bounds the adversary effectiveness more than the key agreement performances. Using a key length of 20 bits, there is no way for the adversary to guess the key but our algorithm performs the agreement in more than 95% of cases.

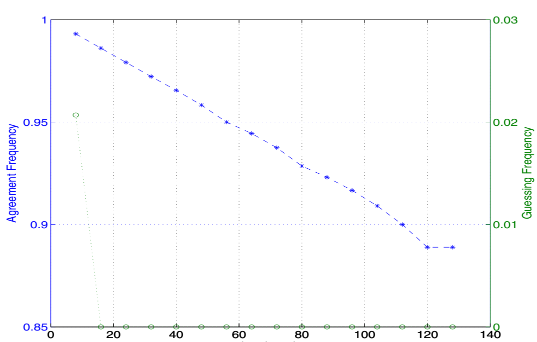
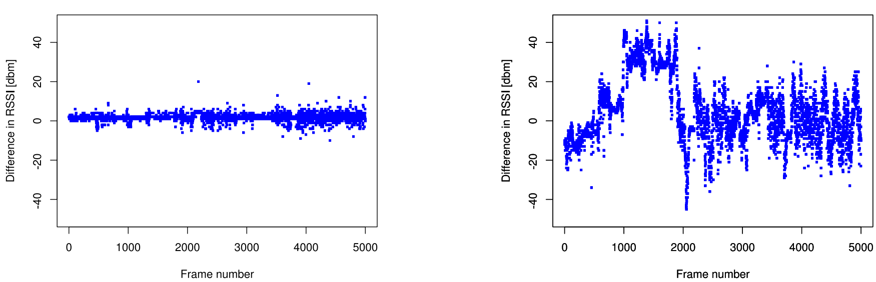


Fig. 6. Key agreement frequency and key guessing frequency as function of the key length, fixing b_w to 27

5.3 How Secret Are the Secret Keys?

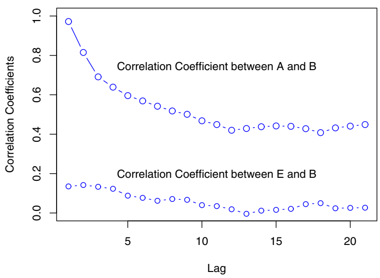
In this subsection we are interested in analyzing the amount of uncertainty that the generated keys contain, i.e., how much information can an adversary $\mathcal{E}$ gather by monitoring the packets exchanged between $\mathcal{A}$ and $\mathcal{B}$. For this reason, we estimate the information (Shannon) entropy of generated symbols, which is defined as $H(K) = -\sum_{i=1}^{n} p(S_i) \log p(S_i)$. Here, the secret key K is a discrete random variable with possible values of quantified RSSI measurements $\{S_1, \ldots, S_n\}$ as described in the previous sections, and p is the probability mass function of K. To estimate the information entropy of generated keys, we first analyze the similarity of the RSSI values measured by $\mathcal{A}$ and by $\mathcal{E}$. A sample trace of 5000 measured RSSI values over both channels is shown in Fig. 7. These results demonstrate how the measured RSSI values are strongly position-dependent. Specifically, while the difference in the RSSI values of the messages exchanged between $\mathcal{A}$ and $\mathcal{B}^0$ are highly correlated and similar (shown by a small difference between RSSI values in Fig. 7 (a)), the same messages captured by the adversary $\mathcal{E}$ result in significantly different RSSI measurements, see, Fig. 7 (b). This insight is also validated by the correlation coefficients, which indicate that the channel of $\mathcal{A}$ and $\mathcal{B}^0$ is highly correlated (≥ 0.6) especially for the first few packets (key derivation uses a lag of 1, additional lags are shown for completeness), while the same packets captured at $\mathcal{E}$ have a negligible correlation coefficient with respect to those experienced by $\mathcal{A}$ (≤ 0.2). Consequently, even if $\mathcal{E}$ can capture the entire communication between $\mathcal{A}$ and $\mathcal{B}^0$, as long as she has a different physical position, her measurements cannot help her to reveal RSSI estimation at $\mathcal{A}$ and $\mathcal{B}^0$. Yet, before we can estimate the information entropy of generated keys, we need to find out the underlying distribution p of symbols $\{S_1, \ldots, S_n\}$. Clearly, the information entropy of K is the highest if p is uniformly distributed. However, this is not the case since RSSI values follow the Rayleigh or Rician distribution depending on the Line-of-Sight component [8]. For this reason, we used the experimental testbed to systematically collected over 10000 generated symbols and analyzed



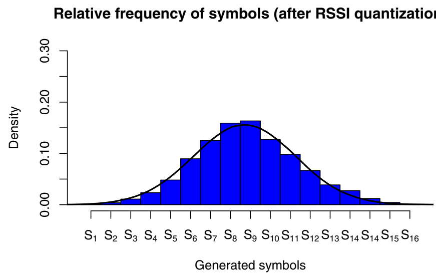
(a) Differences of the measured RSSI values over the $\mathcal{A}$ and $\mathcal{B}^0$ channel.

(b) Differences of the measured RSSI values over the $\mathcal{B}^0$ and $\mathcal{E}$ channel.

Fig. 7. An example of the spatial-dependency of RSSI values: comparing differences in RSSI values from the same packets received at the two different physical positions



(a) The RSSI correlation analysis.



(b) Distribution of the generated symbols used for key derivation.

Fig. 8. Eavesdropping from another physical position on the message exchange between $\mathcal{A}$ and $\mathcal{B}$ results in a highly decorrelated RSSI values (left). The relative frequency of the generated secret symbols follows a normal distribution (right).

their relative frequencies. This analysis reveals a normal distribution of generated symbols as shown in Fig. 8(b) (as an example, we selected the parameter $b_w = 2$, although any other value for b_w would not alter the distribution). Finally, assuming p to be normally distributed random variable $N(\mu, \sigma^2)$ with the parameters estimated using the maximum likelihood method over the sample $\{S_1, \ldots, S_n\}$, we can compute the information entropy of the generated key K. Depending on the b_w parameter, the entropy of generated keys is summarized in the following table and compared with the "best-case" assumption of uniform distribution (see Table 2).

Table 2. Entropy per symbol: normally versus uniformly distributed symbols

b_w	2	4	8	16
Normal Distribution	3.39	2.92	1.57	0.98
Uniform Distribution	4	3.16	2.32	1.58

6 Conclusions

In this paper we propose a new algorithm for securing smart spaces by generating shared secret keys between the user and the AmI infrastructure. The shared secret is generated from the received signal strength experienced by both the user's sensor and the anchors. We justified that the signal power has a symmetric behavior by means of an extensive measurement campaign, moreover, we leveraged it to extract shared secrets. We showed how our algorithm bounds the adversary behavior in four different cases: brute force on the on-fly packet, sensor compromise, RSSI guessing attack, and finally, symbols statistical analysis. In summary, this paper demonstrates how the nature of wireless communication offers valuable properties for security designs and assists AmI applications to minimize the need of user interactions.

Acknowledgments

P. Barsocchi and G. Oligeri's research was partially supported by the EU 7th framework program, contract no. 247950 (universAAL open platform and reference Specification for Ambient Assited Living).

I. Martinovic's research was funded by the Carl-Zeiss Foundation Fellowship.

References

1. Aono, T., Higuchi, K., Taromaru, M., Ohira, T., Sasaoka, H.: Experiments of IEEE 802.15.4 ESPARSKey (Encryption Scheme Parasite Array Radiator Secret Key) - RSSI Interleaving Scheme. In: IEICE Tech. Rep., Kyoto, vol. 105, pp. 31–36 (April 2005)
2. Azimi-Sadjadi, B., Kiayias, A., Mercado, A., Yener, B.: Robust Key Generation from Signal Envelopes in Wireless Networks. In: CCS 2007: Proceedings of the 14th ACM Conference on Computer and Communications Security, pp. 401–410. ACM, New York (2007)
3. Baronti, P., Pillai, P., Chook, V.W.C., Chessa, S., Gotta, A., Hu, Y.F.: Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards. Computer Communications 30(7), 1655–1695 (2007)
4. Crossbow, http://www.xbow.com
5. Girolami, M., Lenzi, S., Furfari, F., Chessa, S.: Sail: A sensor abstraction and integration layer for context awareness. In: EUROMICRO-SEAA, pp. 374–381 (2008)
6. Jana, S., Premnath, S.N., Clark, M., Kasera, S.K., Patwari, N., Krishnamurthy, S.V.: On the effectiveness of secret key extraction from wireless signal strength in real environments. In: MobiCom 2009: Proceedings of the 15th Annual International Conference on Mobile Computing and Networking, pp. 321–332. ACM, New York (2009)
7. Mathur, S., Trappe, W., Mandayam, N., Ye, C., Reznik, A.: Radio-telepathy: Extracting a Secret Key from an Unauthenticated Wireless Channel. In: MobiCom 2008: Proceedings of the 14th ACM International Conference on Mobile Computing and Networking, pp. 128–139. ACM, New York (2008)
8. Rappaport, T.: Wireless Communications: Principles and Practice. Prentice Hall PTR, Upper Saddle River (2001)
9. Wilhelm, M., Martinovic, I., Schmitt, J.B.: Secret keys from entangled sensor motes: implementation and analysis. In: Proceedings of the Third ACM Conference on Wireless Network Security (WiSec), pp. 139–144 (March 2010)
10. Wilson, R., Tse, D., Scholtz, R.A.: Channel Identification: Secret Sharing using Reciprocity in Ultrawideband Channels. In: ICUWB 2007: IEEE International Conference on Ultra-Wideband, pp. 270–275 (September 2007)
11. Zeng, K., Wu, D., Chan, A., Mahapatra, P.: Exploiting Multiple-Antenna Diversity for Shared Secret Key Generation in Wireless Networks. In: Proceedings of the 29th IEEE Conference on Computer Communications (INFOCOM), San Diego, CA (March 2010)

Taxi-Aware Map: Identifying and Predicting Vacant Taxis in the City

Santi Phithakkitnukoon[1], Marco Veloso[2,3], Carlos Bento[2], Assaf Biderman[1], and Carlo Ratti[1]

[1] SENSE*able* City Lab, Massachusetts Institute of Technology, Cambridge, MA, USA
[2] Departamento de Engenharia Informática, Universidade de Coimbra, Portugal
[3] Escola Superior de Tecnologia e Gestão de Oliveira do Hospital, Instituto Politécnico de Coimbra, Portugal
santi@mit.edu, mveloso@dei.uc.pt, bento@dei.uc.pt, abider@mit.edu, ratti@mit.edu

Abstract. Knowing where vacant taxis are and will be at a given time and location helps the users in daily planning and scheduling, as well as the taxi service providers in dispatching. In this paper, we present a predictive model for the number of vacant taxis in a given area based on time of the day, day of the week, and weather condition. The history is used to build the prior probability distributions for our inference engine, which is based on the naïve Bayesian classifier with developed error-based learning algorithm and method for detecting adequacy of historical data using mutual information. Based on 150 taxis in Lisbon, Portugal, we are able to predict for each hour with the overall error rate of 0.8 taxis per 1x1 km^2 area.

1 Introduction

We envision a map-based service platform that allows access to real-time information about the state of taxi transportation as well as predictions regarding its future state. A pilot service that turns available taxi-GPS data into useful contextual information that will be provided to citizens for making taxi transportation more efficient and pleasant to use and to policy-makers as a decision-support tool. As an initial step towards building such a platform, we present in this paper a framework for predicting the number of vacant taxis in a given area. The knowledge of the current state of the taxis (number of vacant taxis) in different areas in the city as well as the future state provides the information for a better scheduling. For example, a tourist who arrives at an airport in a transit city and wants to make a trip inside the city with limited time will benefit from the service by using it to plan out a series of taxi rides around the city. A taxi service provider can also use the platform to monitor their taxis and optimize dispatch scheduling. The envisaged platform paves the way for *Ambient Intelligence* (AmI)'s smart city concept that refers to physical environments in which information and communication technologies and sensor systems disappear as they become embedded into physical objects and the surroundings in which we live, travel, and work [1].

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 86–95, 2010.

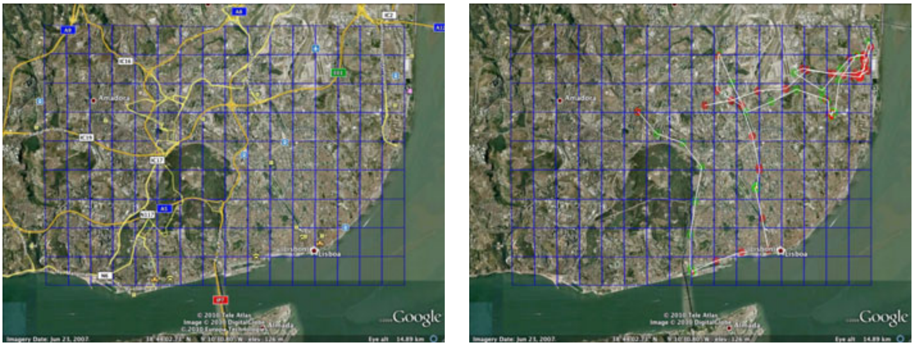
2 Related Work

In recent years, a massive increase in the volume of records of when and where people are has been produced with large deployment of pervasive technologies in the cities. These digital footprints of individual mobility pattern have motivated an increasing number of research in human mobility such as city dynamics [2], predictability of human mobility [3], event-driven traveling pattern [4], and activity-based mobility pattern [5].

GPS-enabled vehicle data such as taxi traces have been collected and analyzed. Chang et al. [6] describe a model that predicts taxi demand distribution based on time, weather condition, and location. Results are shown with different clustering techniques and based on five taxi drivers over two months in service. Yamamoto et al. [7] propose an adaptive routing algorithm using fuzzy clustering to improve taxi dispatching by which vacant taxi drivers are adaptively assigned to pathways with many potential customers expected. Ziebart et al. [8] describe a framework that probabilistically models a distribution over all behaviors (i.e., sequences of actions) using the principle of maximum entropy within the framework of inverse reinforcement learning. With 25 taxis, they show that their model outperform the existing models in turn, route, and destination prediction. Liu et al. [9] analyze 3,000 taxi drivers' operation behavior and categorize them into top and ordinary drivers based on the income. The result reveals the top driver's intelligence through path choice and location.

3 Data Preparation

In this research, we use anonymous data of taxi-GPS enabled traces collected during the period from Sept. 1 thru Dec. 15, 2009 by Geotaxi[10]. This includes 2.6 million anonymous locations of 150 taxis in Lisbon, Portugal. Lisbon is the capital of Portugal with its urban area expanding around the downtown with greatest population density, touristic, historic and commercial areas, and the center for public transportation services. Residential area, airport, and industrial facilities are located in the citys periphery. The data sampling rate varies according to the trip – distance driven, time elapsed, or state changed (e.g. occupancy). Each of the 150 taxi traces carries the information about the taxi's location, service status (occupied, vacant), and the corresponding time and date. If $S = \{s_1, s_2, ...\}$ represents a trace of a taxi, then each instance sample k contains location, service status, and timestamp, i.e. s_k = (latitude$_k$, longitude$_k$, service$_k$, unix timestamp$_k$). For our analysis, we model the map of Lisbon with one-kilometer square grid cells. For each cell, the ID is assigned in raster scanning fashion where it begins with ID #1 at the bottom left corner and increases horizontally left to right. Once it reaches the most right cell, then the counting continues on the next line (upward). The total number of cells in this study is 144 with 16 horizontal cells and 9 vertical cells, as shown in Fig 1(a). To give the readers a sense about our data, Fig. 1(b) shows a sample taxi trace of 10 hours in service, where each dot represents the recorded location with red and green color denoting the service status of being occupied and vacant, respectively.



(a) Area of study (b) A sample taxi trace

Fig. 1. Map of Lisbon is modeled by 1x1 km^2 grids and a sample of a taxi trace during 10 hours in service. The red dot represents reported location when the taxi is occupied while the green one shows when the taxi is vacant.

4 Predictive Model

The objective is to predict the number of vacant taxis in a grid cell for a given time. To do so, in a probabilistic approach, we need an *inference engine* that estimates the probability of some numbers of vacant taxis within a cell given some observables drawn from historical data. The prediction can then be derived according to the maximum probability criterion. Clearly, if the predictor performed perfectly, one would expect no error. In reality, that is not the case. To improve the performance of the predictor, the predictor thus needs to learn from the errors. Hence *error-based learning* is an essential part of the predictor. One of the most critical elements in context-aware computing and AmI system design is the large data processing as the system must deal with the increasing amount of sensory data to build a priori knowledge for inferring the context of the users. Similarly, in our case, we need a method to detect the *adequacy of historical data*, which will reduce the amount of data in repository and computational cost while retaining the amount of information contained in the original data. In this section, we will describe our approach in building the predictive model that consists of inference engine, error-based learning, and adequacy of data detection. The system overview is shown in Fig. 2.

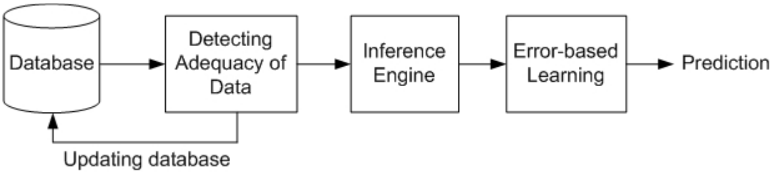


Fig. 2. System overview

4.1 Inference Engine

Taxi drivers prefer the areas with the most potential customers as well as the areas with potential long-trip customers e.g. airport. A taxi, once occupied by a customer(s) will move to the customer's destination via some path (not necessarily the shortest path). Once a customer ride is completed, the taxi becomes vacant and cruises either in the same area or goes to other areas to seek its next customers. Typically, each taxi attempts to minimize the search time for the next customers. It is intuitive that a taxi driver wants to make the most money in the least amount of time, and hopes to pick up many passengers whose destinations are in places where there are customers to perpetuate a chaining of constant business. As random as it seems, taxis can be predictable to some extent. Their mobility patterns are driven by customers and their individual driving strategies (to maximize revenue). Where and when they become vacant is of our interest in this study. Thus, by considering each grid cell individually, a priori knowledge about the number of vacant taxis of given time can be derived from historical data.

As people normally travel according to the business hours around the city, taxis thus move accordingly to meet the demand e.g. more vacant taxis in residential areas in the morning and in the commercial areas in the afternoon. *Time of the day* therefore becomes a reasonable indicator for estimating vacant taxis in an given area. Besides time of the day, *day of the week* seems to carry some clue about the vacant taxis as well. For example, there tends to be more taxis in business areas during the weekdays than in the weekends. *Weather condition* also plays an important role in traveling decision making. A nice sunny day may attract more people to go out and travel compared to a rainy day. The number of vacant taxis thereby varies with the characteristic of the given area captured by a series of observations based on these factors.

Our inference engine is constructed with a *naïve Bayesian classifier*, which is a probabilistic classifier based on Bayes theorem with independence assumptions [11]. In our case, we want to compute the likelihood of each possible number of vacant taxis (Y) given time (hour) of the day (T), day of the week (D), and weather condition (W). The probability of the number of vacant taxis is y_i can be computed as

$$P(Y = y_i|T, D, W) = \frac{P(Y = y_i)P(T, D, W|Y = y_i)}{P(T, D, W)}, \tag{1}$$

where $T = \{1, 2, 3, \ldots, 24\}$, $D = \{\text{Monday, Tuesday, \ldots, Sunday}\}$, and $W = \{\text{Sunny, Cloudy, Rainy}\}$. The prediction is then made using MAP method [11], which selects the number of vacant taxis that (y_{MAP}) that maximizes *a posteriori* as follows:

$$\begin{aligned} y_{MAP} &= \arg\max_{y_i \in Y} P(Y = y_i|T, D, W) \\ &= \arg\max_{y_i \in Y} P(Y = y_i)P(T, D, W|Y = y_i) \\ &= \arg\max_{y_i \in Y} P(Y = y_i)\prod_i P(T|Y = y_i)P(D|Y = y_i)P(W|Y = y_i). \end{aligned} \tag{2}$$

As an example, Fig. 3 shows distributions of vacant taxi volume given different time of the day, day of the week, and weather condition. This sample is based on the history of grid #28. The information of weather condition is obtained from Weather Underground [12].

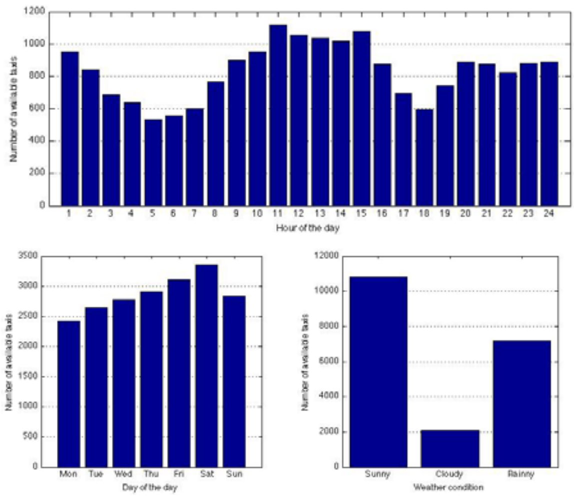


Fig. 3. A sample distribution of vacant taxis given time of the day, day of the week, and weather condition. This sample is drawn from grid #28.

4.2 Error-Based Learning

With the inevitable uncertainty that any predictive model must deal with, the actual outcome may not occur as predicted. Hence error occurs in the system. Instead of letting the error occurs while using it as part of evaluation for the model, we utilize the error to improve the performance of our predictor. Especially the errors that take place most recently indicates the possibly change in pattern. An efficient system must be able to automatically detect and adapt into the change. So the key is to learn from the "recent" errors as the recent data reflects the current trend. Our approach is to apply a weight function that emphasizes on the recent errors from which the prediction is then adjusted accordingly. One possible weight function is the *uniform* function as given by Eq. (3).

$$w_u(t) = \begin{cases} \frac{1}{\beta} & t_m - \beta \leq t \leq t_m \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where t_m is the most recent time and β is the bandwidth (indicating how much of history being taken into account). The weight is distributed equally over the most recent β errors. Another possibility is to assign more weight for more recent errors within the bandwidth. This can be a *linear* function such as the one given by Eq. (4).

$$w_l(t) = \begin{cases} \frac{1}{\beta}(t - t_m + \beta) & t_m - \beta + 1 \leq t \leq t_m \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The prediction will therefore be adjusted according to the weight function as:

$$\hat{y}(t) := \hat{y}(t) + \sum_{i=t-\beta}^{t-1} w(i)e(i), \quad (5)$$

where $\hat{y}(t)$ is the prediction made at time t, $w(i)$ is the weight function, and $e(i) = y(i) - \hat{y}(i)$ is the error measured by the difference between the actual $(y(i))$ and predicted value.

4.3 Adequacy of Data

With the vision of the AmI, which refers to the environment equipped with surrounding computing devices that are sensitive and responsive to the presence and context of people, designers of such environment must deal with the processing of a large human behavioral data. These data can be collected from multiple sensors and processed in some sophisticated way to extract some knowledge about the users. The larger data becomes, the higher capacity is required for storage as well as computation. If the goal is to capture the core structure of some pattern in the data, then the entire data might not be necessary needed. The interesting question is then *how much of historical data is actually adequate to characterize behavioral pattern of interest?*

In our case, the historical data is used to construct a priori distributions for the predictor. Thus, the amount of historical data that allows us to capture the same distributions as we were to use the entire data is "adequate". The behavioral pattern of interest in our case is therefore the distribution of vacant taxis over the grid cells.

To detect the adequacy of historical data, we apply information theory's mutual information [13], which is a measure of the amount of information that one random variable contains about another random variable. Let X denote a random variable representing entire data and Z be a random variable representing some amount of the most recent data in X. The mutual information $I(X;Z)$ is defined as the reduction in the uncertainty of X due to the knowledge of Z, as follows:

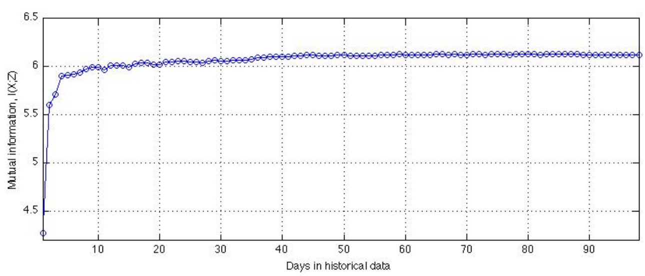
$$\begin{aligned} I(X;Z) &= H(X) - H(X|Z) \\ &= H(X) + H(Z) - H(X,Z), \end{aligned} \quad (6)$$

where $H(X)$ and $H(X,Z)$ are information entropy (uncertainty) of X and joint entropy given by Eq. (7) and (8), respectively.

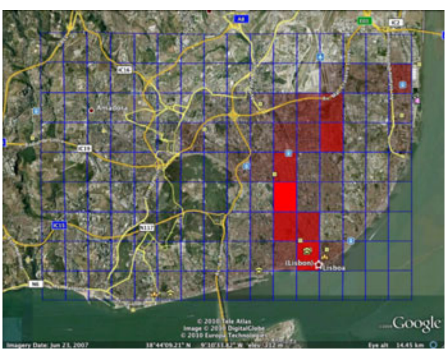
$$H(X) = -\sum_x p(x) log_2 p(x) \quad (7)$$

$$H(X, Z) = -\sum_{x,z} p(x, z) log_2 p(x|z) = -\sum_{x,z} p(x, z) log_2 \frac{p(x, z)}{p(z)} \quad (8)$$

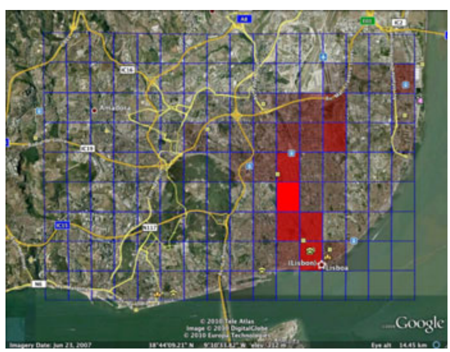
The idea is to find the amount of data carried by Z that can describe most majority of information contained in X. In our case, $p(x_i) = n_i / \sum_i n_i$ where n_i is the amount of vacant taxis in cell i. Based on our experiment, in turns out that a certain amount of historical data carried by Z is adequate to characterize X. As an example, Fig. 4(a) shows the value of $I(X; Z)$ as Z continues to carry more historical data (in days). The value of $I(X; Z)$ converges around 40 days of historical data. Since $I(X; X)$ is equal to $H(X)$ (self-information), the result implies that $I(X; Z) \approx I(X; X) = H(X)$ when at least the last 40 days of data has been taken into account. In the other words, the last 40 days of data is adequate to describe the entire data. Figure 4(b) and (c) show similar distributions of vacancy derived from the entire data and the last 40 days, respectively. This example is one of the taxi traces in our dataset.



(a) Mutual information value that converges as the number of days in historical data increases



(b) Distribution of entire data



(c) Distribution of retained data before convergence (40 days)

Fig. 4. A sample from Taxi ID #9 of convergence of mutual information values and distributions of entire data as well as retained data of recent 40 days. The color shade represents the value of the distribution over different grids.

5 Experimental Results

To test the performance of our predictive model, the first 30 days of data is used for training the model while the rest of the data (77 days) is used as a testing set. A prediction of the number of vacant taxis is made for the next 24 hours (one for each hour slot – 24 predictions are made for each testing day) for each grid cell. The errors are then computed and used to adjust the later predictions (as described in Sect. 4.2). The testing data will sequentially become training data after each prediction has been made. Adequacy of historical data is detected each day of testing (as described in Sect. 4.3) and the training set is then updated accordingly. To detect the convergence time, we use a method described by Phithakkitnukoon and Dantu [14] with threshold of 0.02.

During the experiment, the absolute error is computed as the absolute difference between the actual and predicted amount of vacant taxis, i.e. $|\hat{y} - y|$. The cumulative error is computed over all grid cells and shown in Fig. 5. To show the improvement of the predictor achieved by the error-based learning and detection of adequacy of data, in Fig. 5, cumulative errors of the model with (full model) and without these features (baseline model) are shown, where clear improvement of about 10% in error can be observed. The uniform and linear weight functions are used (with $\beta = 5$) where both functions yield similar performance. The overall error per grid cell is shown in Fig. 6 where cells in blue, yellow, and red have error range of 0–1, 1–2, and 2–3 taxis, respectively. The higher error cells seem to be clustered in commercial and touristic areas while lower error cells are spreading gradually. Overall, the errors are in the range from zero to three taxis, which is considerably promising. In the other words, the reliability of each prediction is $\pm c$ where $c = 1$, 2, and 3 if the cell is blue, yellow, and red, accordingly.

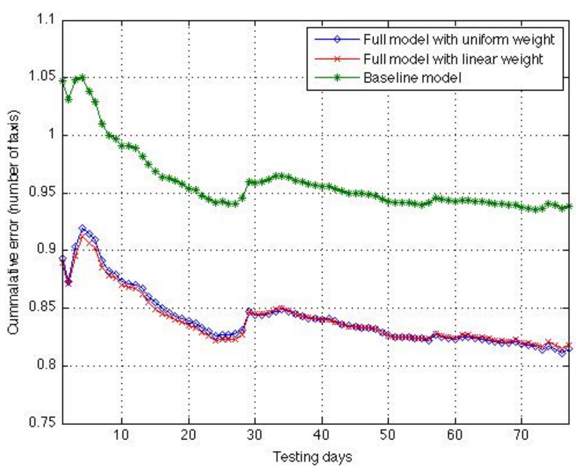


Fig. 5. Cumulative error over days in testing period. The full model with the error adjustment and adequacy of data outperforms the baseline model.

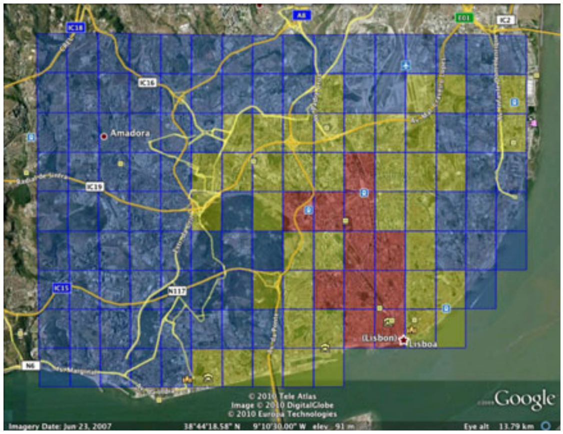


Fig. 6. Error rate in different grid cell. Error ranges are 0–1, 1–2, and 2–3 taxis, which are represented with blue, yellow, and red respectively.

6 Conclusions

With the vision of smart cities, which is an application of concept and development in AmI, here we take the first and essential steps toward building a map-based service platform that allows access to real-time information about the state of taxi transportation as well as predictions regarding its future state. This will be useful for taxi service providers, transportation management, as well as taxi users. In this paper, we develop a framework for predicting the number of vacant taxis in an area based on time of the day, day of the week, and weather condition. Our predictive model is based on Bayesian classifier with a sequential error-based learning algorithm and a mechanism for detecting the adequacy of historical data using information theory's mutual information. Based on 150 taxis in Lisbon, Portugal, our model is able to predict (per hour) accurately with the overall error of slightly over 0.8 taxis per 1x1 km^2 area. As our future direction, we will continue to find ways to improve our framework such as by using different weight functions (e.g. exponential, sigmoid), weight bandwidth, and exploiting multi-source data fusion (combing data from different sources e.g. mobile phone, bus, fleet). A larger dataset will also be obtained (from Geotaxi) for our future development.

Nonetheless there is a number of limitations of this study that we are aware of as following:

- The sample size of 150 taxis may not be the true representative of the whole taxi population. It would be interesting to see how our framework performs with a larger sample size.
- The grid size is relatively large and may not be realistic for individual users who are particularly interested in the prediction within a smaller area. In this study, the grid size is designed to compensate the taxi sample size that

we have. With a larger sample size, smaller grid size can be implemented with the same developed framework. Nevertheless, the current grid size is still useful for taxi service management and urban/transportation planning.
– The weather condition is assumed to be the same throughout the day. For example, if the record shows that September 1st, 2009 was a sunny day, then the weather condition is assigned to be sunny for entire 24 hours, which may not be true. Nonetheless, we do not anticipate a large change in the overall result if we had accounted for the change of weather condition during day.

Acknowledgments. The authors gratefully acknowledge support by Volkswagen of America Electronic Research Lab, the AT&T Foundation, the National Science Foundation, MIT Portugal, and all the SENSEable City Laboratory Consortium members.

References

1. Steventon, A., Wright, S.: Intelligent Spaces: The Application of Pervasive ICT (Computer Communications and Networks). Springer, New York (2005)
2. Reades, J., Calabrese, F., Sevtsuk, A., Ratti, C.: Cellular census: Explorations in urban data collection. IEEE Pervasive Computing 6(3), 30–38 (2007)
3. Song, C., Qu, Z., Blumm, N., Barabsi, A.L.: Limits of predictability in human mobility. Science 327(5968), 1018–1021 (2010)
4. Calabrese, F., Pereira, F.C., Lorenzo, G.D., Liu, L.: The geography of taste: analyzing cell-phone mobility and social events. In: Proceedings of IEEE Inter. Conf. on Pervasive Computing (PerComp) (2010)
5. Phithakkitnukoon, S., Horanont, T., Lorenzo, G.D., Shibasaki, R., Ratti, C.: Activity-aware map: Identifying human daily activity pattern using mobile phone data. In: Inter. Conf. on Pattern Recognition (ICPR 2010), Workshop on Human Behavior Understanding (HBU), pp. 14–25. Springer, Heidelberg (2010)
6. Chang, H., Tai, Y., Hsu, J.Y.: Context-aware taxi demand hotspots prediction. Int. J. Bus. Intell. Data Min. 5(1), 3–18 (2010)
7. Yamamoto, K., Uesugi, K., Watanabe, T.: Adaptive routing of multiple taxis by mutual exchange of pathways. Int. J. Knowl. Eng. Soft Data Paradigm. 2(1), 57–69 (2010)
8. Ziebart, B.D., Maas, A.L., Dey, A.K., Bagnell, J.A.: Navigate like a cabbie: probabilistic reasoning from observed context-aware behavior. In: UbiComp 2008: Proceedings of the 10th International Conference on Ubiquitous Computing, pp. 322–331. ACM, New York (2008)
9. Liu, L., Andris, C., Bidderman, A., Ratti, C.: Revealing taxi drivers mobility intelligence through his trace. In: Movement-Aware Applications for Sustainable Mobility: Technologies and Approaches, pp. 105–120 (2010)
10. Geotaxi, http://www.geotaxi.com
11. Mitchell, T.M.: Machine Learning. McGraw-Hill, New York (1997)
12. WeatherUnderground, http://www.wunderground.com/
13. Cover, T.M., Thomas, J.A.: Elements of information theory. Wiley-Interscience, New York (1991)
14. Phithakkitnukoon, S., Dantu, R.: Adequacy of data for characterizing caller behavior. In: Proceedings of KDD Inter. Workshop on Social Network Mining and Analysis (SNAKDD 2008) (2008)

Dynamic Privacy Management in Pervasive Sensor Networks

Nan-Wei Gong, Mathew Laibowitz, and Joseph A. Paradiso

Responsive Environments Group, MIT Media Laboratory
E14-548, 75 Amherst Street, Cambridge 02142, USA
{nanwei,mat,joep}@media.mit.edu

Abstract. This paper describes the design and implementation of a dynamic privacy management system aimed at enabling tangible privacy control and feedback in a pervasive sensor network. Our work began with the development of a potentially invasive sensor network (with high resolution video, audio, and motion tracking capabilities) featuring different interactive applications that created incentive for accepting this network as an extension of people's daily social space. A user study was then conducted to evaluate several privacy management approaches – an active badge system for both online and on-site control, on/off power switches for physically disabling the hardware, and touch screen input control. Results from a user study indicated that an active badge for on-site privacy control is the most preferable method among all provided options. We present a set of results that yield insight into the privacy/benefit tradeoff from various sensing capabilities in pervasive sensor networks and how privacy settings and user behavior relate in these environments.

Keywords: dynamic privacy management, ubiquitous computing, active badge system, pervasive sensor networks.

1 Introduction

As we move into the era of Ambient Intelligence, we will see a shift in how people interact with information. Just as we are now noting users evolve into often using mobile devices instead of desktops and laptops, we will see a shift into everywhere interaction with pervasive smart environments when sensing, actuation, and display are embedded ubiquitously into our surroundings, and our digital "cloud" manifests on whatever devices and information portals are available and appropriate. The user interface to this environment will accordingly also be abstracted into a ubiquitous sensor network that acts as the perceptive "nervous system" of ambient intelligence. The sea of sensors that surround us are already mushrooming, as we bring more and more sensing into our presence on the back of devices we acquire for specific services. Once common standards enable applications to share sensor data across devices, we will see an explosion of development similar to what happened when the Web united networked servers and PCs. It's vital, however, that before we reach this point, the sensors rushing into in our lives respond intrinsically to our dynamic desires for privacy [1,2] – there will just be too many to manually turn off or disable.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 96–106, 2010.

Accordingly, this paper relates a set of studies that we have run at our laboratory to determine how users accommodate such a large pervasive multimodal sensor/display network and gain insight into how their behavior adapts to a set of dynamic privacy protocols that we have developed.

We began the study of privacy in ubiquitous interactive sensor networks with the installation of 45 "Ubiquitous Media Portals" (UMPs) in our building. This potentially invasive sensor network is equipped with high-resolution video, audio and motion tracking capabilities. The UMPs ran several applications [2-4] to engage users to interact with those sensing affordances. The video and audio captured by each sensor node can be streamed between different nodes for image sharing and message broadcasting, as well as to online platforms such as Second Life for ubiquitous virtual-reality applications [3]. We also developed several wearable sensors that augment this network with on-body sensing to assemble meaningful content, such as a user-generated documentary video[4-5].The displays on each UMP can show information of general interest, such as the latest RSS feeds. They also allow users to capture and share images or text messages that are broadcast to all the other nodes throughout the building. During an8-month pilot study before we constructed the privacy management system, we implemented four different applications to explore acceptance by our building-wide users. The results verified that applications allowing sufficient, transparent interaction and providing generally useful information are effective ways to increase the percentage of nodes remaining active without being physically disabled by our users [2].Learning from this experience, we subsequently built a dynamic privacy management system on the UMPs. Our privacy system consisted of two parts: onsite privacy control (with beacons from active wearable devices, physical switches, and touch screen inputs) and remote privacy settings (via web browsers for setting pre-established privacy preferences).

1.1 Previous Work

Privacy Protocol Design. Substantial research has been devoted to design strategies and policies for privacy issues in ubiquitous computing environments. The major approach for controlling privacy status within sensor networks is through constructing secure protocols and code verification mechanisms for system developers to follow and examine as they deploy the infrastructure for data acquisition and post data processing. Bellotti and Sellen were pioneers with their work on privacy in the context of ambient video based on the experience of the RAVE media space at EuroPARC. They first proposed a privacy-protected framework in 1993 [6] for designing ubiquitous computing environments and described the ideal state of affairs with respect to each of four types of behavior – Capture, Construction, Accessibility and Purposes. Their argument is that providing obvious "feedback and control" over information in a ubiquitous computing environment can help assuage privacy concerns (Sellen and colleagues have recently demonstrated a system of networked cameras and pen-enabled displays for interhousehold interaction in Microsoft's Wayve system [7]). Drawing upon their work, many toolkits have been developed to provide programming support and abstractions for protecting privacy in a ubiquitous computing environment such as Confab [8] and Mist [9].

Context-Aware Systems. Researchers have explored protecting privacy through pseudonyms, dummy users, and storing privacy information as a watermark to blur users' information, especially location-specific data, in computer vision [10-12, 13]. Recently, research into privacy protection in context-aware pervasive systems has advanced to the design of self-configuring privacy management infrastructures. Ortmann et al. proposed a self-configuring privacy management architecture for pervasive systems [14]. Further, Moncrieff and coworkers [15] presented a dynamic method for altering the level of privacy in the environment based on inferred context and local situation. Beyond research on dynamic privacy configuration that exploits fixed sensor infrastructure, the concept of automatically inferring mobile privacy settings is also explored through monitoring the use of personal electronic devices such as cell phones [16-17]. All of the above examples demonstrate the idea of creating a smarter and sophisticated system that could better suit users' needs of privacy within their environment. However, without direct user control, the construction of an ideal system that can suit everyone's needs is almost impossible.

Sensing Type and Location Control. Another major method for improving the design of privacy protection in sensor networks is through the physical approach — the privacy-concious choice of sensors and location/direction of sensing elements. In [18], Reynolds and Wren examined the ethical implications of choosing camera networks vs. infrared motion detector networks. Their results indicate that for most participants, infrared sensors were significantly less invasive than pan-tilt-zoom cameras. This comes at the cost, however, of not implanting sensors that can facilitate more complex applications. Although it has been proven that data collected from motion sensing can indirectly lead to approximate personnel identification and localization [19], the coarse level of interaction provided by a motion sensor network still can not yield all functions increasingly desired in evolving ubiquitous networks. Therefore, we try not to compromise our sensor system design, but rather to control the quality of the data provided according to the dynamic privacy level requested by the users' privacy management settings.

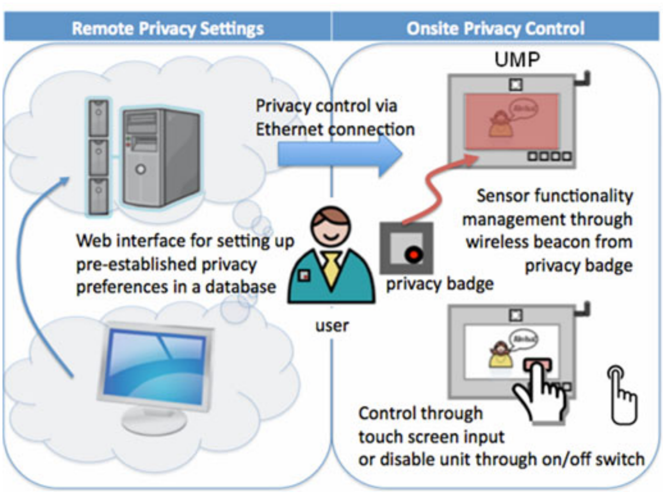


Fig. 1. System block diagram

2 Design and Implementation

The goal of our system is to construct a user-centric privacy management system for ubiquitous computing and use it to obtain real-world experience with a potentially invasive pervasive sensor network. There are two major parts in our privacy management system –onsite privacy control and remote privacy settings (Fig 1).

Our sensor network can communicate with the wearable privacy badges through a building-wide 802.15.4 ZigBee radio network allowing the badges to change the sensing parameters onsite, i.e., turn on or off different sensors according to the settings prespecified by each individual badge user (Sec. 2.2). Also, local users can physically disable the sensing units through touch screen inputs or an on/off power switch.

On the other hand, users can set up their privacy preference online from a web interface by setting the allowable sensor modalities that can stream from each node when they are nearby. Their privacy level can also be dependent on the group status of the client browsing the sensor network—the badge user can assign different levels of privacy to different groups of people (e.g. taking an analogy to UNIX file system permission: “user/group/world”), i.e., individuals who are socially closer to the user can be allowed to have more access. Physical means of providing immediate privacy are also afforded (e.g., physically obstructing the sensors and turning the obviously-located power switch off).

2.1 Active Privacy Badges and Ubiquitous Media Portals

In order to provide active control, we designed and implemented active privacy badges for the UMPs and demonstrated the possibility of integrating the wearable electronics into every day accessories. Fig 2(a) shows the badge we used for our user study. This simple wearable sensor has 4 LED indicators to display users’ dynamic personal privacy settings, a blocking button for requesting immediate privacy, and a vibration motor to alert upon unexpected events, such as a user’s privacy setting being over-written by another nearby user with high priority. It communicates with our UMP sensor network using 802.15.4, through which we have attained room-level RF location accuracy, which is sufficient to disable any UMPs that are in sensing range of the badge wearer. Subsequent badges that we have developed for other projects have added a small display (Fig 2(b) [7,8]) that can alert people near the user to their privacy level or data streaming status.

The need for a privacy system was inspired by our pervasively-deployed multimodal sensor and display Ubiquitous Media Portal (UMP, see figure 3) network. The 45 UMPs that we built comprise a sensor network that was distributed throughout the Media Lab. Each UMP, mounted on a pan/tilt platform, has an array of sensors, as well as audio and video capabilities (see Figure 3 for the list of features). Video and images are acquired with a 3 MegaPixel camera above a touch screen display. The video board is driven by a TI DaVinci processor (an ARM9 running Linux paired with a C64x+ DSP core for video processing), and features a touch-screen LCD display, LED floodlight, &speaker. The sensors and an 802.15.4 radio that talks to and tracks wearable sensors are mounted on a daughter card, which runs

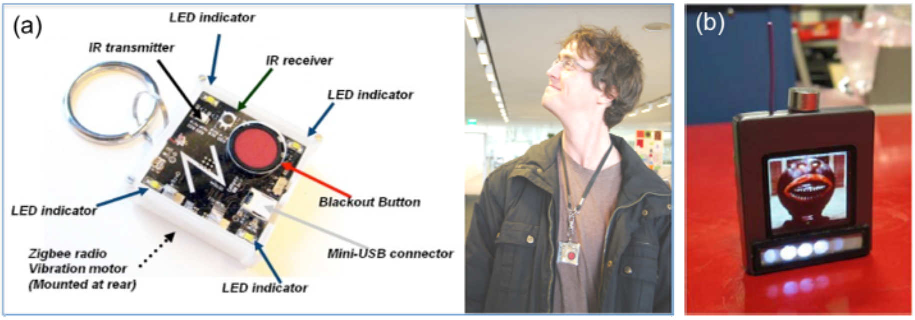
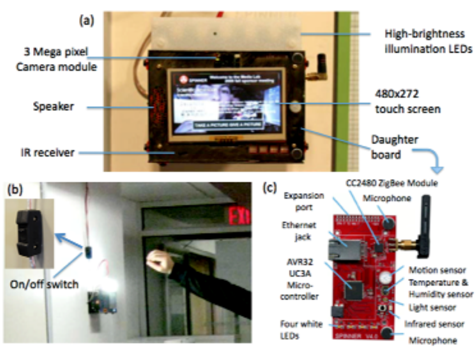


Fig. 2. (a) Privacy Badge – diagram and worn. (b) Subsequent badge featuring OLED display.

an AVR32 microcomputer (AV32UC3A1256) and features stereo microphones, PIR motion sensor, humidity/temperature sensor, light sensor, and 2 protocols of IR communication (for detecting active badges within the line of sight) [7].

2.2 Data Server and Web Interface

For personalized remote privacy control, users can register on the web interface with their unique badge ID and edit their privacy preferences on each node. Although the work by Beresford and Stajano [13] indicated that in a restricted space, pseudonymous ID could be cracked by the location information of each badge, our active badge can disable forwarding of tracking information by proximate UMPs if the user wants to remain anonymous.



Sensors	I/O Channels
3M pixel auto focus camera with pan and tilt control and 5 Watt LED illuminator	480x272 LCD touch screen
Stereo microphones	Speaker
Passive Infrared Sensor	4 LEDs for status report
Temperature and humidity sensor (SHT15)	Zigbee 802.15.4 transceiver
Light sensor	IR communication

Fig. 3. Left: (a) Overview of a UMP node and one example application ("Cloud of Images") showing RSS feeds and the last captured image. (b) UMP, in action illuminating subject for video capture, and switch for manually deactivating UMP's power. (c) Sensor daughter board. Right: List of features in the sensing system.

Users can edit privacy/sensor preferences on a location basis via a web form. For example, in Fig 4, the user is editing the behavior of node 311 and turning off the video recording when they are present at this location. An "edit all sensors" page was also provided to specify globally common settings. Results from how people use this page can give us insight into how our system is perceived [20]– e.g., whether specific locations or the nature of different sensors is the greatest privacy threat for most people. In all of our experiments, the default setting for all sensors was on (making an opt-out system).

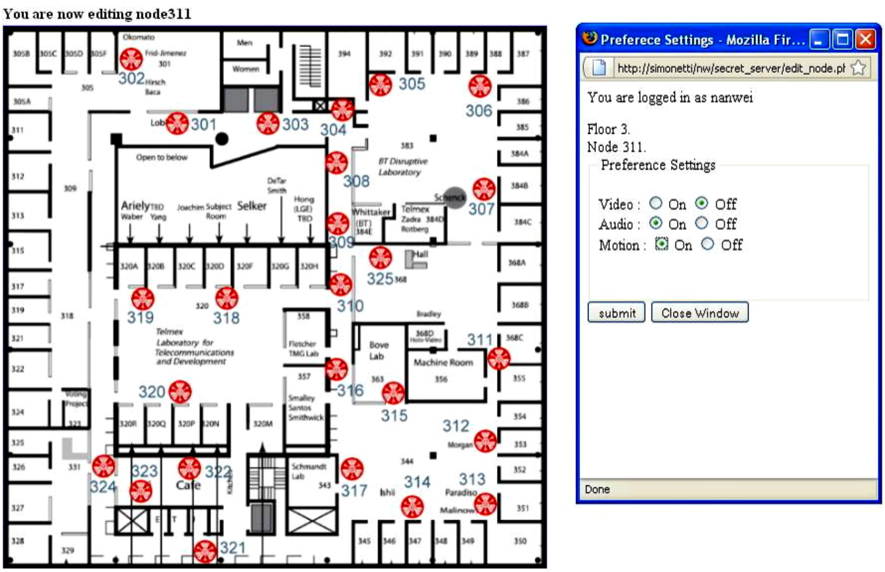


Fig. 4. The "edit sensor" page allows users to click on each node on an interactive map and edit sensor settings on a location basis.

One of the most important aspects about users' privacy protection in a ubiquitous computing sensor network is having the ability to post-process our personal information flow. While our system has the ability to collect video, audio and images and display that information recorded for each user individually, it could also be tailored to share users' information with others. In the edit group permission page, the users are allowed to reveal their information according to social hierarchy -- user / group / world, like a UNIX file permission system (e.g. family, friends, and world in real life). Further, users are able to create their own group and send out invitations for other users to join their group. This framework can not only allow the users to customize how they appear to who is looking, but also can be used as a social networking tool similar to Google Tracker, which let you follow your friends' or families' locations and show ubiquitously-collected images or videos in real-time.

3 Evaluation

We conducted several user studies to gain insight into dynamic privacy management. The first study included four different applications with different levels of control and interaction on the users' end. Users could enable or disable the UMPs with the physical power switch onsite. Our results, collected across an 8-month study, indicate that applications allowing sufficient, transparent interaction and providing generally useful information are effective ways to increase the percentage of nodes remaining on. Detailed application and statistics information can be found in [3]. With this experience, we conducted another user study that featured our active badge system. Users had three different ways to disable sensor streaming— disabling UMPs by cutting their power, using the privacy badge's "NO" button to immediately block data transmission, and modifying their online settings to automatically disable sensor streaming at specific locations when nearby. The on/off switch will kill the entire UMP until it is powered/booted again and rejoins the system (taking many minutes), while the privacy button and the online preset behavior only disable the UMPs temporarily. In these tests, the default blackout time instigated by the privacy button is 10 seconds, which is approximately the typical time needed to walk by a UMP node or sufficient time to blank minimal personal identifying information during a conversation.

Twenty-four people (out of 90 people working on the same floor where all portals were installed) volunteered to participate in this weeklong study. The recruitment requirements [20] include selecting people whose offices are located near the UMPs and obtaining a diverse group of people, including students, staff and faculty members in our building (65% had engineering/science background – the remainder had background in arts, humanities, or other expertise). All were well aware of the privacy threats that our system posed, and knew that video/audio could be streamed from each portal.

3.1 User Study Results

Results from Active Badge Usage. Fig. 5 (left) shows the normalized distribution of on-site "NO" button presses versus time. The users generally press the NO button on their badge when there is a conflict between the location-specified privacy of their online settings and the location of an unexpected private event, or when unquestioned privacy is immediately needed. During our one-week user study, the average number of button presses per day was 71;3 button presses per user per day. The peak correlates with afternoon break / lunch (2-3 PM) with broad tail late into the afternoon when more social interaction is typically exercised and users tend to desire more confidentiality.

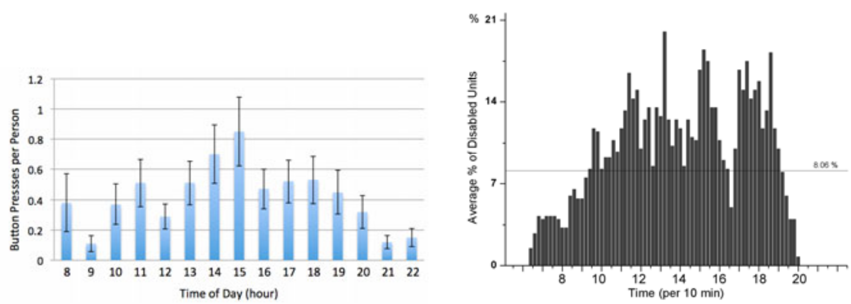


Fig. 5. Left: Ratio of NO button presses per the number of users (error bars are derived from the number of people per bin). Right: Normalized histogram of disabled UMPs over time (the average fraction of disabled UMPs is 8%).

Although we built the active badge system to solve the privacy issue for dynamic situations, gaining a balance between maintaining the full function of a ubiquitous interactive sensor network while preserving users' privacy in every way is still an unsolved issue. Figure 5 (right) shows the average fraction of disabled UMP units from both button presses and automatic shuttering (determined by user proximity to the UMP and their web-assigned preferences) per 10-minute interval. We see up to 20% of the units disabled by privacy requests at peak social intervals.

Results from Users' Online Settings. From the web database, we observed that 70 percent of participants set up their privacy preferences online and 66 percent of all participants chose to block all video transmission through active badge proximity (50 percent set all audio off), whereas only 34 percent set all motion sensors off. As for the privacy settings on individual nodes, 50 percent of all participants set up individual privacy on a location basis. On their daily route, 33 percent turned off the video recording/broadcasting, another 17 percent disabled both video and audio, but merely 4 percent disabled the motion sensors. The remaining 8 percent did not set up preferences from this page, and just used their on-badge button.

Figure 6 explores the relationship between different privacy management methods versus the location of individual UMP units. This plot is segmented according to the location of each UMP. Nodes in group A are in the corners and hallways where people pass by, generally with less social interaction. These nodes are on most of our users' routing path; however, less people set up intrinsic privacy control there online. Group B are nodes in common areas such as our café and kitchen, where most social interactions take place – accordingly, these places encountered both significant online scripted blocking and spontaneous button presses. The last group, C, are nodes located in office clusters and around intersections between different paths. Fewer users marked those places as high privacy risks online, but lots of unexpected private events happened as we can see from the high percentage of privacy button presses.

Results from Questionnaire and Interview. After the evaluation period, each user was asked to complete a post-experiment questionnaire that asked about the usability of this system and the acceptance of each privacy protection method. When users were asked about which approaches could better suit their need for privacy control, 40%answered "online privacy settings", 33%answered "on-site badge control" and another 19% and 8%answered "button on the touch screen" and "cutoff switch" respectively.

We also interviewed the users and ask them to list the scenarios in which they used the active badge system for privacy incidents. 75 percent of users listed "to block audio recording from an unexpected private conversation"; 50 percent listed "to block the video recording". One interesting finding is that, among the users who listed "video blocking" as one of the irrationales, some mentioned that they prefer not to be recorded when they are alone in a public area, such as in front of the elevator or in the kitchen. This result supports the idea that privacy is a dynamic factor and cannot be generalized easily.

Also, 21 percent indicated that they pressed the NO button to block location tracking occasionally. Though it is possible to block all sensor data collection, many people decided to keep the motion/location tracking on and block the more invasive signals via the privacy badge– these people tried to create a participatory "location presence" in front of their colleagues. On the other hand, 12.5 percent of users never used the badge and commented that they didn't care about being recorded or tracked in their workspace. This indicated that it is impossible to define a standard privacy scenario. Only a personalized system with dynamic controlling capabilities can provide a privacy-enhanced ubiquitous media environment. Our active badge system satisfied this for most of our users.

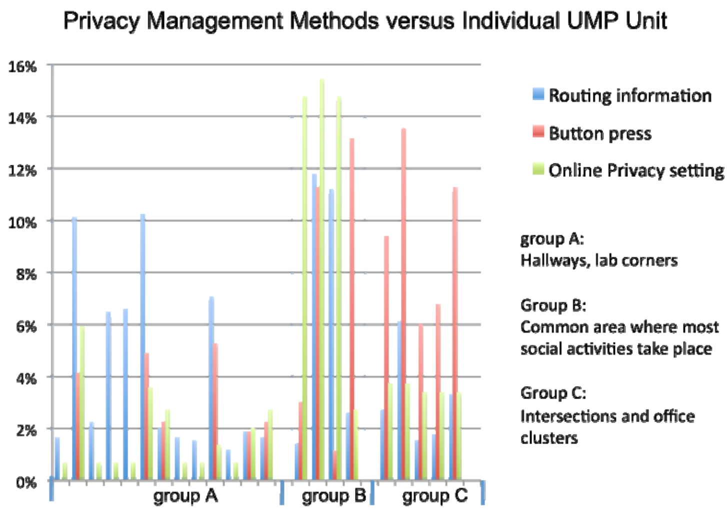


Fig. 6. Relationship between privacy management methods versus individual UMP location. The "Routing Information" shows how much of the user's specified daily route through the buiding passes near that portal –"Button Presses" are the recorded NO events at that location, and the last quantity shows the percentage of people who chose online to disable that portal when they were nearby.

4 Conclusions and Future Work

In this paper, we presented multiple approaches for personal privacy management in ubiquitous sensor networks. The major contribution to privacy research in ubiquitous computing environments is providing a user-centric privacy-protected platform and obtaining experiences with the deployment of a potentially invasive sensor network.

We have evaluated the usability of an active privacy badge system and the possibility of using this system as a building-wide privacy protection facility. Our results indicated that an active badge system for privacy control is the most acceptable method among all the tested options (disabling data transmission from an active badge system, on/off switches, or the touch screen displays). The results from these tests also suggested that if occupants of moderately denser buildings block data transmission in their vicinity at the rates we see now, the availability of the sensor network will be compromised. Therefore, it is crucial to find a balance between protecting privacy and maintaining enough data flow for the value-added applications utilizing the network at the same time. Ongoing research by various teams (e.g., [21]) is exploring ways of removing particular individuals from audio/video and sensor streams when privacy is desired, maintaining network function. Our future work will explore this avenue, as well as integrating our privacy badge functionality into commodity cell phones and common short-range radio standards like Bluetooth.

This project focused more on interfaces for privacy management and control rather than network security – the protocols and systems that we used were minimally secured, which wouldn't be permissible in an actual deployment. For this type of system to be really used, trusted, and ultimately accepted, communication protocols, software, and hardware need to be implemented that are secure and resistant to attack [22] so that security and privacy can be certifiably protected according to individual choice. More detailed information is available, including all the collected data, higher resolution images, source code, and schematics, in the associated graduate thesis [20].

Acknowledgments

We thank our colleagues at the Media Laboratory who participated in the user study and the deployment of this system, especially Bo Morgan for developing the SPINNERD server. This work was supported by the Things That Think Consortium and the other research sponsors of the MIT Media Laboratory.

References

1. Davies, N., Gellersen, H.-W.: Beyond Prototypes: Challenges in Deploying Ubiquitous Systems. IEEE Pervasive Computing 1(1), 26–35 (2002)
2. Gong, N.-W., Laibowitz, M., Paradiso, J.A.: Experiences and Challenges in Deploying Potentially Invasive Sensor Systems for Ubiquitous VR Applications. In: Cloud-Mobile Convergence for Virtual Reality Workshop (CMCVR 2010), Waltham, MA, March 20 (2010)
3. Lifton, J., Laibowitz, M., Harry, D., Gong, N.W., Mittal, M., Paradiso, J.A.: Metaphor and Manifestation – Cross Reality with Ubiquitous Sensor/Actuator Networks. IEEE Pervasive Computing Magazine 8(3), 24–33 (2009)
4. Laibowitz, M., Gong, N.-W., Paradiso, J.A.: Wearable sensing for dynamic management of dense ubiquitous media. In: Proceedings of 6th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2009), Berkeley, CA, June 3-9, pp. 3–8 (2009)
5. Laibowitz, M., Gong, N.-W., Paradiso, J.A.: Multimedia Content Creation using Societal-Scale Ubiquitous Camera Networks and Human-Centric Wearable Sensing. In: Proc. of ACM Multimedia 2010, Florence Italy, October 25-29 (2010)
6. Bellotti, V., Sellen, A.: Design for privacy in ubiquitous computing environments. In: Proceedings of the 3rd European Conference on Computer-Supported Cooperative Work (ECSCW 1993), pp. 77–92. Kluwer Academic Publishers, Norwell (1993)
7. Lindley, S.E., Harper, R., Sellen, A.: Design of a Technological Playground: A Field Study of the Emergence of Play in Household Messaging. In: Proc. of CHI 2010, Atlanta, GA, April 10-15, pp. 2351–2360 (2010)
8. Hong, J.I., Landay, J.A.: An Architecture for Privacy-Sensitive Ubiquitous Computing. In: Proceedings of MobiSys 2004, Boston, MA, June 6-9, pp. 177–189 (2004)

9. Campbell, R., Al-Muhtadi, J., Naldurg, P., Sampemane1, G., Mickunas, M.D.: Towards security and privacy for pervasive computing. In: Proceedings of International Symposium on Software Security, Tokyo, Japan, November 8-10, pp. 1–15 (2002)
10. Zhang, W., Cheung, S.-C.S., Chen, M.: Hiding privacy information in video surveillance system. In: Proc. of ICIP 2005, Genova, Italy, September 11-14, vol. II, pp. 868–871 (2005)
11. Schiff, J., Meingast, M., Mulligan, D., Sastry, S., Goldberg, K.: Respectful cameras: Detecting visual markers in real-time to address privacy concerns. In: International Conference on Intelligent Robots and Systems (IROS), October 29-November 2, pp. 971–978 (2007)
12. Kindberg, T., Fox, A.: System Software for Ubiquitous Computing. IEEE Pervasive Computing 1(1), 70–81 (2002)
13. Beresford, A., Stajano, F.: Location Privacy in Pervasive Computing. IEEE Pervasive Computing 2(1), 46–55 (2003)
14. Ortmann, S., Langendörfer, P., Maaser, M.: A Self-Configuring Privacy Management Architecture for Pervasive Systems. In: Proc. of the 5th ACM International Workshop on Mobility Management and Wireless Access (MOBIWAC), Chania, Crete Island, Greece, October 22, pp. 184–187 (2007)
15. Moncrieff, S., Venkatesh, S., Andwest, G.: Dynamic privacy in a smart house Environment. In: Proceedings of the IEEE International Conference on Multimedia and Expo, July 2-5, pp. 2034–2037. IEEE Computer Society, Los Alamitos (2007)
16. Gisch, M., De Luca, A., Blanchebarbe, M.: The Privacy Badge - A Privacy-Awareness User Interface for Small Devices. In: ACM International Conference On Mobile Technology, Applications, And Systems, Singapore, September 10-12, pp. 583–586 (2007)
17. Sadeh, N., Hong, J., Cranor, L., Fette, I., Kelley, P., Prabaker, M., Rao, J.: Understanding and capturing people's privacy policies in a mobile social networking application. Personal and Ubiquitous Computing 13(6), 401–412 (2009)
18. Reynolds, C.J., Wren, C.R.: Worse is better for ambient sensing. Workshop on Privacy Trust and Identity Issues for Ambient Intelligence (Pervasive 2006), Technical Report TR2006-005, Mitsubishi Electric Research Laboratories (MERL) (May 2006)
19. Beresford, A.R., Stajano, F.: Location Privacy in Pervasive Computing. IEEE Pervasive Computing 2(1), 46–55 (2003)
20. Gong, N.-W.: Configurable Dynamic Privacy for Pervasive Sensor Networks, MS thesis, MIT Media Laboratory (2009)
21. Wickramasuriya, J., Alhazzazi, M., Datt, M., Mehrotra, S., Venkatasubramanian, N.: Privacy-Protecting Video Surveillance. In: Proc. of the SPIE International Symposium on Electronic Imaging (Real-Time Imaging IX), San Jose, CA, February 25, pp. 64–75 (2005)
22. Stajano, F.: Security for Ubiquitous Computing. John Wiley and Sons, New York (2002)

Geo-Social Interaction: Context-Aware Help in Large Scale Public Spaces

Nasim Mahmud[1], Petr Aksenov[1], Ansar-Ul-Haque Yasar[2], Davy Preuveneers[2], Kris Luyten[1], Karin Coninx[1], and Yolande Berbers[2]

[1] Hasselt University - tUL - IBBT,
Expertise Centre for Digital Media,
Wetenschapspark 2, B-3590 Diepenbeek, Belgium
[2] Department of Computer Science, Katholieke Universteit Leuven,
Celestijnenlaan 200A, B-3001 Leuven, Belgium
{nasim.mahmud,petr.aksenov,kris.luyten,karin.coninx}@uhasselt.be,
{ansarulhaque.yasar,davy.preuveneers,yolande.berbers}@cs.kuleuven.be

Abstract. We present an approach to exploit social and spatio-temporal context in order to improve information dissemination in dynamic large-scale public spaces. We illustrate it by applying a proposed measure of geo-social relevance of each individual in a simulated vehicular network and by comparing the performance of different network message passing techniques in an inter-vehicle 'help-me-best-and-do-it-fast' communication scenario. We conclude that the use of social networking capabilities of an individual combined with knowledge about their spatio-temporal context information significantly improves purposeful interaction between individuals in terms of both the efficiency of the network data dissemination and the quality of the delivered information.

Keywords: Social network, FOAF, Context awareness, Location awareness, Information dissemination.

1 Introduction

Being able to obtain the right information at the right time has always been a challenge when taking informed decisions. Unfortunately, people who are 'on-the-move' often do not have an opportunity to spend a long time looking for what they need. Systems (e.g., intelligent transportation systems) also may require efficient delivery of timely and relevant information, e.g. to optimise traffic flows. The proliferation of novel wireless network technologies has created new opportunities for complex peer-to-peer information dissemination systems, and a key challenge in this area is how to interact, locate and communicate effectively in a large scale public environment.

Context-awareness is often an instrument to decide what information is relevant and as such can help to improve communication and routing efficiency since it allows making informed decisions on the locality and information necessity of different moving objects. In our previous work, we have shown that context

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 107–116, 2010.

awareness improves network efficiency in a vehicular network by exploiting the user's context in making informed decisions on the data routing [7]. In order to provide improved help and facilitate users to achieve their goals, the system should be able to recognise their context, such as the current place, state, activity, interests, etc. Among the many facets of context, location has been widely recognised as one of the primary context areas and it is the main concern of many modern mobile applications. The social context has also recently taken one of the leading roles in context-awareness. The importance of these two areas have, in turn, led to the concept of location-based social networks and services. Here, social and spatiotemporal branches of context are closely tied together, having given rise to a new type of 'geo-social' interaction, collaboration and information sharing. Some recent examples of the work in this area include the Connected Traveler [4] project that aims at providing relevant information to drivers by means of taking into account their quickly changing location and personalised preferences in the way that only the information the driver explicitly set to be interested in is selected. Connecto [1] is a phone based status and location sharing application, which evolved, throughout use within a small group of friends, from serving for mere location updates into a tool for enriched social interaction through these location-based updates. CityFlocks [2] is a mobile system that allows information seeking visitors to access tacit knowledge from local people about their new community. Commercial applications, such as Foursquare[1] and Gowalla[2], allow sharing one's current location and recommending (or ranking) favourite places to friends, which proved to be a very attractive addition to the basic social networking capabilities.

These and many other examples show that the area of using location and social context together is very broad. In this paper, we investigate how exploiting a social network, user's personal preferences, information about user's location and movements as well as presence and preferences of other users allows for interaction that helps users in large-scale highly dynamic and populated public spaces to obtain timely and relevant information. In particular, we aim at selecting the most appropriate individual as an assistant in a 'Help me!'-like scenario between members of a vehicular network. Our approach involves a comparison of social and spatial relevance between two peers in terms of the area of expertise and spatial closeness. We have utilised social network analysis to find the most appropriate peer in terms of the possessed knowledge and the distance to cross and have factorised the urgency of the help seeker in terms of time the assistance is being looked for.

2 Three-Leaved Mirror Approach

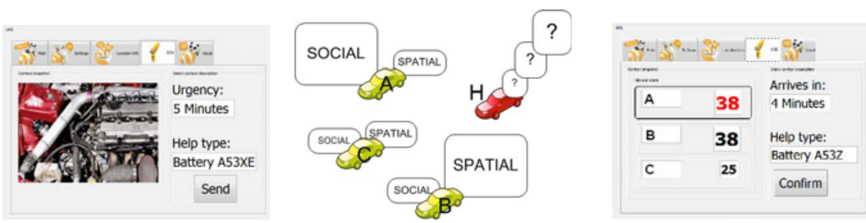
In our three-leaved mirror approach, the spatial and the social contexts each resemble one of the two side leaves of such a mirror. Like it is with a real three-leaved mirror when the angles of the leaves can be adjusted independently

[1] Foursquare: http://foursquare.com
[2] Gowalla: http://gowalla.com

to change the current view of the user from each side, the contribution of the spatial and social context to the overall relevance varies. The central leaf then optimises the information flow in the network by reducing irrelevant information dissemination, thus making the view of the user on the issue at hand more complete, enabling to see 'the big picture'.

2.1 The Side Leaves: Geo-Social Component

Like network availability, location-awareness has become a ubiquitous property of many mobile devices nowadays. With both networking and location detection technologies readily available in the device itself, we can combine the information available on the user's location with information provided by the user's social network. For example, we can easily select a contact from the user's social network that is in the vicinity and ask this person for assistance. The difficulty lies in dealing with the additional complexity caused by persons in the social network moving and the timeliness of certain requests for assistance (e.g. when no help arrives within 5 minutes the requester will have to manage on their own). We used vehicular networks for evaluating our approach. Since this type of network is highly dynamic, contains fast moving objects and requires a close eye on the timeliness, it serves as the perfect evaluation framework to assess our approach.



(a) (b) (c)

Fig. 1. Talking cars. 1(a) Help seeker H sends a help request using an embedded Geo-social Ubiquitous-Help-System (UHS); 1(b) A, B, and C are ready to help H; 1(c) Depending on the returned geo-social relevance score of the helpers, H confirms the offer.

2.2 The Central Leaf: Improved Relevance Backpropagation with Geo-Social Relevance

Geo-social interactions require communication between people, and optimised information dissemination is its vital aspect. There exist a number of information routing strategies, both with and without taking into account the quality of information (QoI) (e.g., broadcasting and backporpagation, respectively). For example, Eichler et al. [3] address the issue of optimal information dissemination in vehicular networks by proposing a framework which integrates many of the existing broadcast-based strategies that deal with reduction of the superfluous transmissions. We propose a QoI-based best-effort mechanism for intelligent adaptive context dissemination using social interactions with a relevance score function (see section 4) in a large scale vehicular network based on our previous work [5]. In general, the technique can be applied to many types of networks where social interactions matter, and currently we chose vehicular networks since they are a comfortable means for verification as they contain sufficient diversity in terms of both size and dynamics. In our case, each participant has a list of friends (e.g., in a Friend-of-a-friend (FOAF) profile), a score value and properties about the relevant context information they can provide. The algorithm relies on the feedback of neighbours to reduce the number of peers to forward the information to. The information is forwarded to the adjacent nodes who are either friends or friends-of-a-friend having certain degree of score value unless a maximum number of hops is reached. Each forwarding node reduces the hop counter, adds its identifier and marks the message relevancy tag if the information is relevant for its purpose and grades the sending node positively adding it to the friends list. The feedback technique is based on the context information like position, velocity, direction, time-to-live, interest, etc., that decides whether the received data are relevant which helps determine the information relevancy on the intermediate nodes. The feedback to the delivering node is initiated if the context information is relevant, irrelevant, unused, or duplicate information is received. It ensures that the provided information is from a trusted node which is supposed to be accurate and relevant for the receiver. In this mechanism the goal is to efficiently filter and route the relevant information as close to the source as possible in vehicular ad-hoc networks (VANETs).

3 Aspects of Geo-Social Interaction

Our three-leaved mirror approach poses a set of generic aspects to consider for providing context-aware help in large scale public spaces.

Spatial coverage. It is always desirable to know the exact location of an incident for context-aware applications. for instance, in the case of an incident on the road, the authorities should be notified about the exact location to react fast and the information should be delivered only at the right place. Similarly, a context-aware application should be able to sense, manipulate and disseminate context information about direction and velocity of vehicles in the network to predict certain situations like traffic congestions or traffic accidents in specific regions.

Timeliness. It is crucial that the information being disseminated in a large scale network between nodes (or groups of nodes) reaches the destination on time. Timeliness uses time as a relevance criterion for information sharing so that the information received by a node is not older than the 'lifetime of the information' so that the right information can be delivered at the right time.

Completeness. Lack of information can lead to ambiguity. This measure identifies the quality of information (QoI) that is provided by a node in the network and can be comparing the number of attributes received to the total number of attributes to make a well-informed decision.

Trust-worthiness. The information being received at a particular node should be reliable, accurate and trustworthy. Each node involved in a large scale vehicular network should have a social profile listing sharing some common interest with a certain quality of information (QoI).

Significance. It indicates the importance of a certain type of contextual information required by a node in a network. The value of significance will be higher in the case of a life threatening situation.

4 The Geo-Social Relevance Function

The key concept to our approach is the geo-social relevance function that defines how relevant a potential help provider is to the help seeker. It naturally accumulates together certain parameters from each of the three leaves (spatial, social, and network-bound) and results in a sort of weight each involved node of the network has with respect to the help requester. The higher the score, the better the provider. The function is expressed using the following formula:

$$GSR(peer) = A \cdot \sqrt[n]{R \cdot HT} \cdot F_U$$

where the multipliers in the right part are as follows: **Availability A** is a Boolean value that simply indicates whether the corresponding node can be a potential help provider. We assume that if a peer is unavailable ($A = 0$), the help request is still passed further to this peer's friend-list. **Reliability R** of a node in the network is a peer-determined integer value between 1 and 10 indicating how helpful the corresponding node has been in the past. **Help-type HT** measures the requester's and the provider's technical match and is an integer between 0 ('I know nothing') and 10 ('A perfect match'). **Root index n** stands for the number of hops in the network between the requester and provider. The reason for choosing the root-based value for measuring the contribution of the social parameters to the overall relevance is that the level of trust to somebody who is connected to you indirectly decreases significantly. **The spatiotemporal contribution,** F_U, is defined as

$$F_U = \begin{cases} 0 & \text{if } Direction = 0 \\ 0 & \text{if } U \cdot V \leq D_{min} \\ e^C & \text{if } D_{min} \leq U \cdot V \leq D_{max} \\ e \cdot C & \text{if } U \cdot V > D_{max} \end{cases}, \text{ where } C = \frac{U \cdot V - D_{min}}{D_{max} - D_{min}}$$

Here, ***Direction*** equals 1 if there is a 'movingTowards' relation from the provider to the requester, and 0 otherwise. Urgency ***U*** is the time interval within which the help is needed; its value is specified by the requester. Velocity ***V*** is an estimated average velocity of the help provider during period U. We assume that the help requester does not move (e.g., their car broke down). The corresponding maximal and minimal distances, D_{max} and D_{min}, between two nodes are calculated at the time when the help request has been received and depend on the spatial topology of the area, such as, for example, the actual length of the connecting path between the nodes that might be affected by possible repair works, closed or blocked roads, etc. The reasoning behind the expression for F_U is such that for nodes that are far $(0 \leq C < 1)$, $e^C > e \cdot C$ and thus e^C has a bigger weight for the score; and for $C \geq 1$, the smaller value of $e \cdot C$ is used thus making the social parameters weigh more than the spatial ones when comparing the scores of two different nodes.

This way, there exists a sort of balance between the social relativity and the spatial closeness so that nodes with different social profiles, but in the direct vicinity, can get equal, or close enough, scores and be chosen by the help requester in accordance with the requester's personal preferences. Besides, the score function makes all members of the network - the requester (urgency U), potential providers (availability A), and all other nodes in the network (reliability R) - collaborate implicitly in finding the fittest solution to the help request.

5 Validation by Simulation

In order to enable intelligent communication between vehicles we need to take into account the social characteristics of the people involved. Since large scale groups are not easily tested with real-life situations, we evaluated our improved relevance backpropagation algorithm using a real time discrete event-based network simulator (OMNeT++)[3] to run on a large scale vehicular network using a realistic dataset [6] logged for a period of 24 hours. Individuals in the simulation are distributed over an area of 250 by 260 km. All individuals choose a time to travel and a route in accordance with where they live and current road congestion. The complete dataset contains more than 25,000,000 recorded direction and speed changes of 260,000 vehicles, from which we randomly selected 300 vehicles. Besides, for the sake of simplicity, we normalised the original movement data to fit in the range of $[0; 1]$.

5.1 Data Preparation and Setup

Taking into account the aspects discussed in section 3, this section describes the details of the actual data we used in the simulation.

Friendlist. In general, the number of friends in the friend-list of a node in the network is not limited. However, given the specifics of our simulation and the actual behaviour the simulation is reproducing, we limited each node to have a maximum of 15 friends, the exact pre-simulation initial number of friends being assigned based on the type of help a node can provide (see Table 2). Friend-lists get extended in an asymmetric way: each time a help requester has received help from an appropriate provider, the latter is added to the requester's friend list. In the case of a new help request from the same node, chances are that the same provider can help again, and contacting this provider directly will save time and resources. An extract from the complete assignment showing the initial friend lists of the first five nodes is given in Table 1.

[3] OMNeT++: `http://www.omnetpp.org`

Helptype. We introduced nine help-types so that each of the 300 nodes belongs to one of them. The total number of nodes distributed normally among each help-type is as follows using the (help-type - number of nodes of this type) notation: (1-11), (2-26), (3-43), (4-55), (5-55), (6-45), (7-32), (8-21), and (9-12).

The matching table of the HT values corresponding to a pair of help-types is shown in Table 2. Notice the asymmetric nature of the HT values for the corresponding (ReqHT1 to PrHT2) and (ReqHT2 to PrHT1) pairs, meaning that in general, 'If you are able to help me with my problem, it does not guarantee that I am able to help you with yours.'

Table 1. Initial distribution of friends into friendlists for the first five nodes

Node	F1	F2	F3	F4	F5	F6	F7	F8	F9	···	F15
1	79	117	93	0	0	0	0	0	0	···	0
2	10	1	54	74	126	165	184	0	0	···	0
3	10	4	77	44	88	84	189	177	0	···	0
4	11	64	186	0	0	0	0	0	0	···	0
5	74	76	113	97	150	273	0	0	0	···	0

Table 2. Pair-matching of help-types

	PrHT1	PrHT2	PrHT3	PrHT4	PrHT5	PrHT6	PrHT7	PrHT8	PrHT9
ReqHT1	10	0	8	8	8	0	0	7	0
ReqHT2	8	10	0	0	0	8	6	8	4
ReqHT3	8	0	10	9	7	0	8	6	8
ReqHT4	9	4	9	10	9	0	4	0	8
ReqHT5	6	6	4	7	10	6	9	8	0
ReqHT6	7	8	0	4	0	10	7	9	9
ReqHT7	4	7	7	8	6	8	10	0	6
ReqHT8	0	9	0	0	4	9	8	10	7
ReqHT9	0	0	6	6	8	7	0	4	10

Reliability. An integer value between 1 and 10 was assigned randomly to each node's reliability. In general, reliability increases or decreases dynamically based on nodes' performance as helpers but for the purposes of our simulation, we kept the help-type values static throughout the simulation.

Urgency. Out of the 300 nodes, we picked 10 which would, each at a random point during the simulation time, become a help requester with an individual urgency value the help request is valid for. The urgencies of 600, 300, 120, 180, 120, 120, 120, 90, 180, 540, and 240 seconds, were used by each help requester, respectively.

Since the available measurements of nodes' movements contained only a timestamp and 2D-coordinates, other required values had to be further derived. Thus, the average velocity of each node over a certain time interval was calculated using the next two subsequent measurements for the node in question which gave an acceptable approximation in terms of the time the help requests remained valid for. There were also no data on the location error of the provided measurements. Therefore we divided the entire area into three sub-areas throughout each of which a measurement had a precision value from a specified interval. The precision of a node's current position is a characteristic of a localisation system, which tracks the node's position, and can be defined as the radius of a circle centered in the detected location. The circle then means the node is located somewhere within it. We set the precision intervals to be $0.0005 - 0.0001$, $0.0005 - 0.001$, and $0.001 - 0.005$, respectively, with 0.0001 equal to approximately 25 metres on the original scale. Since we did not have a description of the spatial topology of the area, we assumed the distance between two nodes to be a simple Euclidian distance. Together with the assigned precisions, it gave the following formulae for computing the D_{min} and D_{max} values between two nodes at the time of calculation:

$$D_{min} = min(0, D_{mes} - P_A - P_B)$$
$$D_{max} = D_{mes} + P_A + P_B$$
$$D_{mes} = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}$$

where P_A and P_B are the precision values of the corresponding measured locations of the two interacting nodes. Notice that in the case when the precision circles intersect, D_{min} equals 0.

5.2 Details of Simulation Runs

In our experiments, we let the nodes move around like cars and let connections appear and disappear according to the range to other nodes. The parameters we have taken into account for each node are: (i) Time, (ii) Velocity, (iii) X and Y coordinates, (iv) Number of packets sent, (v) Number of packets received, (vi) Number of forwarded packets and (vii) Time-to-live(TTL). Some nodes acted as context providers and some as context receivers. All nodes forward the information to their peers as long as the maximum TTL has not been reached and all context constraints are met.

We carried out three experiments with (a) our improved relevance backpropagation mechanism with social interactions and the geo-social relevance function, (b) simple relevance backpropagation and (c) state-of-the-art baseline case, plain broadcasting, each for a period of 24 hours.

6 Results

With both the (improved and simple) relevance backpropagation algorithms only relevant context information was sent out to the interested nodes in the network. There are several types of messages in the network: (i) sent (M_s), (ii) unique received (M_{ur}), (iii) unique sent (M_{us}) (iv) forwarded (M_f), (v) duplicate (M_d), and (vi) dropped (M_{drop}). During our simulated experimentation, we measured a set of major network metrics, Network Traffic (NT), Relevancy (R), Trustworthiness (T), Message Distance (MD), and Availability (A), expressed in terms of the above network messages as follows:

$$NT = \Sigma_n(M_s + M_f)$$
$$R = \Sigma_n((M_{ur} + M_d) - M_{drop})/\Sigma_n(M_{ur} + M_d)$$
$$T = (1 - (M_f/\Sigma_f(n)(M_{ur} + M_d)))$$
$$MD = \Sigma_t Edges/\Sigma_t Nodes$$
$$A = \Sigma n M_{us}/\Sigma n M_{ur}$$

For each involved metric, we achieved a significant improvement in its performance using the proposed improved relevance backpropagation as compared to both simple relevance backpropagation and broadcasting. For example, the utilisation of the Network Traffic is 90% of that using simple relevance backpropagation and as low as 50% of that of simple broadcasting. The five left plots in Figure 2) visualise such comparisons for all five metrics on the percentage scale.

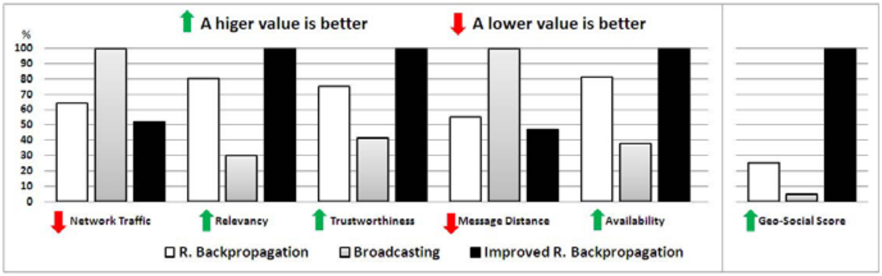


Fig. 2. Improved relevance back propagation technique for routing messages in the network shows better results for each evaluated parameter

We also measured the performance of the geo-social relevance function in three cases of our experimental setup. Some returned scores were quite high in simple relevance backpropagation algorithm (e.g. 96) and low in the improved relevance backpropagation scheme (e.g. 16) for a certain type of help required. This meant that in some cases unexpected or unplanned nodes who just happened to be close enough had a better match than the algorithmically chosen ones, but when averaged over all returned values the improved algorithm outperforms the other (see the right plot in Figure 2).

7 Discussion and Future Work

The simulation results show that our improved relevance backpropagation mechanism achieves a significant improvement in terms of several network quality of information (QoI) parameters like relevancy, message distance, network traffic, availability and trustworthiness as discussed in section 6. The results in Figure 2 show that by eliminating redundant and irrelevant information sources with the help of the geo-social relevance function we can limit the information dissemination to happen within a much smaller number of nodes having a high degree of relevance, reliability and trustworthiness, thus improving the overall performance of a vehicular network.

The dataset we used for simulations contains limited attributes and focuses in the area of vehicular networks. Therefore we plan to conduct a deeper study on the actual values of social and location parameters (such as distribution of individuals' properties, help-type match values, etc.), augment a reward based mobile search in the UHS and later on cross validate our research in different application domains like pedestrian navigation or care-giving in hospitals. We also plan to improve the visualisation of ranking of search results provided to a user by conducting usability tests in a real life setting (e.g. inside a car).

Acknowledgments

Part of the research at EDM is funded by EFRO (European Fund for Regional Development) and the Flemish Government. Part of the research at K.U.Leuven is funded by the Interuniversity Attraction Poles Programme Belgian State, Belgian Science Policy, and by the Research Fund K.U.Leuven. Funding for this research was also provided by the Research Foundation Flanders (F.W.O. Vlaanderen, project CoLaSUE, number G.0439.08N).

References

1. Barkhuus, L., Brown, B., Bell, M., Sherwood, S., Hall, M., Chalmers, M.: From awareness to repartee: sharing location within social groups. In: Czerwinski, M., Lund, A.M., Tan, D.S. (eds.) CHI, pp. 497–506. ACM, New York (2008)
2. Bilandzic, M., Foth, M., De Luca, A.: Cityflocks: designing social navigation for urban mobile information systems. In: DIS 2008: Proceedings of the 7th ACM Conference on Designing Interactive Systems, USA, pp. 174–183 (2008)
3. Eichler, S., Schroth, C., Kosch, T., Strassberger, M.: Strategies for context-adaptive message dissemination in vehicular ad hoc networks. In: Proceedings of the 2nd Int. Workshop on Vehicle-to-Vehicle Communications (2006)
4. Manasseh, C., Ahern, K., Sengupta, R.: The connected traveler: using location and personalization on mobile devices to improve transportation. In: LocWeb. ACM International Conference Proceeding Series, p. 9. ACM, New York (2009)
5. Preuveneers, D., Berbers, Y.: Architectural backpropagation support for managing ambiguous context in smart environments. In: Stephanidis, C. (ed.) UAHCI 2007 (Part II). LNCS, vol. 4555, pp. 178–187. Springer, Heidelberg (2007)
6. Raney, B., Voellmy, A., Cetin, N., Vrtic, M., Nagel, K.: Towards a microscopic traffic simulation of all of Switzerland. In: ICCS 2002: Proceedings of the Int. Conference on Computational Science, London, UK, pp. 371–380 (2002)
7. Yasar, A.-U.-H., Mahmud, N., Preuveneers, D., Luyten, K., Coninx, K., Berbers, Y.: Where people and cars meet: social interactions to improve information sharing in large scale vehicular networks. In: SAC 2010: Proceedings of the 25th ACM Symposium on Applied Computing, USA, pp. 1188–1194 (2010)

The Operator Guide: An Ambient Persuasive Interface in the Factory

Alexander Meschtscherjakov, Wolfgang Reitberger, Florian Pöhr, and Manfred Tscheligi

Christian Doppler Laboratory for "Contextual Interfaces"
HCI & Usability Unit, ICT&S Center
University of Salzburg
Salzburg, Austria
{alexander.meschtscherjakov,wolfgang.reitberger}@sbg.ac.at,
{florian.poehr,manfred.tscheligi}@sbg.ac.at
http://www.icts.sbg.ac.at

Abstract. In this paper we introduce the context of a semiconductor factory as a promising area for the application of innovative interaction approaches. In order to increase efficiency ambient persuasive interfaces, which influence the operators' behaviour to perform in an optimized way, could constitute a potential strategy. We present insights gained from qualitative studies conducted in a specific semiconductor factory and provide a description of typical work processes and already deployed interfaces in this context. These findings informed the design of a prototype of an ambient persuasive interface within this realm - the "Operator Guide". Its overall aim is to improve work efficiency, while still maintaining a minimal error rate. We provide a detailed description of the *Operator Guide* along with an outlook of the next steps within a user-centered design approach.

Keywords: factory, ethnography, UCD, ambient persuasion.

1 Introduction

Researching human computer interaction in the office has had a long tradition, whereas other working environments have been investigated only more recently. From an HCI perspective the factory in general and even more specifically semiconductor factories, have been rather neglected. So far only very few studies have taken place in this context, which is characterized by high complexity, 24/365 productivity and clean rooms. From a semiconductor company's perspective the most relevant goals within the production cycle are on the one hand the improvement of productivity with as little resources as possible. On the other hand a low error rate is essential to provide constantly high product quality and reduce the loss of defective products. Even more a zero-defect production is desired since in semiconductor factories produced goods are used in applications where highest reliability is required (e.g. automotive applications). In order to support operators during their daily work, a variety of interfaces are currently used within

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 117–126, 2010.

the factory. (In a semiconductor factory the term "operator" denotes the factory's workers in the clean room.) However, little work has been done so far to foster coordination and decision support for the operators in a situated manner. Thus, the aim of our work was to include insights gained from the emerging field of ambient persuasion [9] into the design of a novel guiding system for operators - the proposed *Operator Guide*. The *Operator Guide* constitutes an ambient persuasive interface within the semiconductor factory in order to influence the behaviour of operators following a predefined persuasive strategy. Following key AmI paradigms, it is designed to use already built mental models of the operator as well as being unobtrusively integrated into the environment.

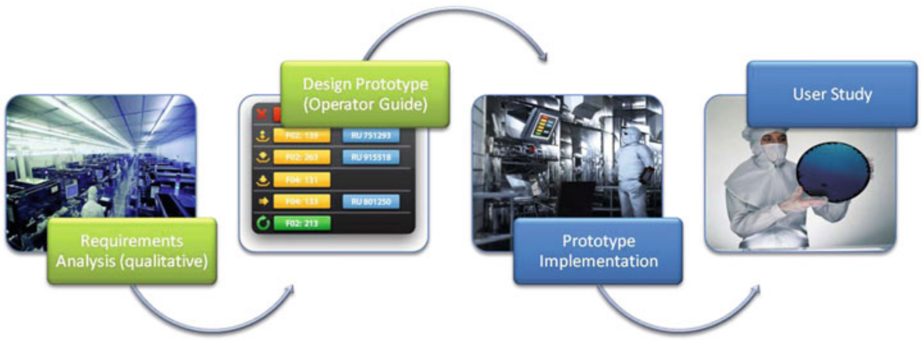


Fig. 1. User-centered design process in the factory

In this paper we present the first two steps of a user-centered design process (see figure 1). Within the user requirement analysis phase we conducted a series of qualitative studies within a specific semiconductor factory. These studies provided us with a holistic understanding of the factory environment in general, basic operator work tasks and already deployed interfaces. Thereafter we used these insights to inform the design of a prototype of the *Operator Guide*. The next steps (an implementation and validation of the *Operator Guide*) will be not part of this paper. After providing an overview of related work we describe the setup of several qualitative studies within a particular semiconductor factory. The following chapter describes the general context of the semiconductor factory as well as a selection of already deployed interfaces. Based on our findings we then present the *Operator Guide* as an initial design approach.

2 Background and Related Work

Ambient Intelligence (AmI) refers to the pervasion of the everyday world with information technology [1]. The users are surrounded by multiple "invisible" systems, which communicate with each other and interact with many users in a wide range of dynamically changing contexts. Persuasive technology is defined as "any interactive computing system designed to change people's attitudes or behaviours" [4]. Ambient persuasion combines these research fields allowing to surround the user with persuasive technology in their everyday life, giving the possibility for persuasive interventions just at the right time and in the right place [9].

Despite the fact that ambient persuasion has been studied in different contexts (e.g. [7]), the environment of a semiconductor factory has not been a target area for this kind of research yet. To get a deep insight into this complex context ethnography constitutes a possible way. Nevertheless, there is "no one method of ethnographic analysis" [6] but a set of methods. This includes, but is not limited to fieldwork, participant observation, interviews, questionnaires, and rather novel approaches like Holtzblatt's contextual inquiry (CI) approach [5] . Within the factory context ethnographic methods have been used in a couple of scientific studies. A case study of maintenance engineering in a steel mill was conducted by Pipek et al. [8] including on-site visits, semi-structured interviews, and participative workshops. Bertelsen and Nielsen [2] presented a study of unskilled work in a Danish wastewater treatment plant. During 6 days within a 5-month period different researchers followed different workers, using hand-held video cameras to capture the events for later analysis. Reul [10] conducted a CI to identify problems with an industrial control system that is used to command and monitor a plant.

Applications and artifacts which support workers within the factory context have been proposed from an HCI perspective only sparely. For instance, Fallman [3] presented the design of a wearable system (PDA-bades, arm-worn, gesture-driven, perceptually seductive and context-aware) that support service technicians with their daily work. Yarin and Ishii [11] presented "TouchCounters" which are physical storage containers with electronic labels for usage frequency and correlation visualization. In conclusion, even though ambient technologies have been researched in various contexts using qualitative methods, and a view design approaches for ambient artifacts within factories have been presented the investigation of interfaces in the clean room through a HCI lens constitutes an interesting and rather novel challenge.

3 Requirement Analysis Setup

To research the factory context we carried out a contextual inquiry (CI) followed by an ethnographic study in a specific semiconductor factory. While the goal of the CI was to get a first glance at the factory context and its specific characteristics, the aim of the ethnographic study was a detailed understanding of the work processes and conditions. Our CI involved three researchers following different operators during their regular shift work. Overall we conducted eight inquiries in different areas of the factory, each CI lasting approximately two hours. Additional interviews to understand organizational aspects were held. Due to the CI we became acquainted with the conditions in a clean room and the restrictions that came along with it. Writing on special paper with special pens, wearing rubber gloves is not convenient. Since the semiconductor industry is a highly competitive area, confidentiality is important. As a consequence videotaping is strictly forbidden. Audio recordings are limited due to background noise and the mandatory face masks of the operators.

The experiences we made in the CI were used to inform the setup of the ethnographic study. One member of our research group worked similar as a trainee in the factory. Prior to the study the co-workers within the shift and the work council were informed about the study setup and goals. Overall the ethnographic study lasted for two shift cycles with six times eight hours each. During the study the scientist worked together with the other operators within a specified production line. This made annotations during the shift nearly impossible. Therefore the researcher was equipped with an audio recorder. He commented his work and experiences by talking to himself holding the recorder close to his mouth. Other operators were not recorded due to the surrounding noise and their face masks. At the end of each shift the researcher wrote down his memories of the day. After completion these annotations were sent to the rest of the research group on a daily bases to provide instant feedback. After the ethnographic study the data was analyzed with qualitative data analysis software. In the next chapter we describe the context of the semiconductor factory gained from our results.

4 Requirement Analysis Results

A semiconductor factory is a complex interplay of several entities. Its overall purpose is to manufacture as many error free integrated circuits as possible. A fundamental step during manufacturing is the processing of wafers, which are thin slices of semiconductor material, such as silicon crystal. Based on the manufacturing algorithm wafers may spend several weeks in the production line. This extensive processing time makes wafers an extremely expensive good. Thus one of the main goals from a company perspective is to avoid errors leading to a "zero-defect-factory" paradigm. Another important goal is - of course - a high productivity. This means to get wafers through the factory as fast as possible with little idle time resulting in an maximized equipment load.

Wafers are typically combined into groups of 25 or 50 pieces, stored in plastic containers called "lot boxes" (see figure 2). Each of these *lot boxes* has to complete a distinct path through the factory, during which it undergoes different process steps (e.g. etching, exposure, etc.) performed on various equipment. Due to different manufacturers and types the interfaces of the equipment are very heterogeneous. To each type of equipment highly specialized operators are assigned. Whenever *lot boxes* need a particular processing step on their well-defined way through the factory operators have to receive them, process them on the right equipment, and deliver them to the next destination. In particular a group of one to four operators work within a specific area which consists of a number of equipment and delivery shelves for incoming and outgoing *lot boxes*. Thus, at any given moment the operator has to assign *lot boxes* to certain equipment. This means that for each *lot box*, the operator has to make at least two decisions. Firstly, he has to decide whether he processes this specific *lot box* now

Fig. 2. Left: Clean room in a semiconductor factory; Center: Operator moving *lot boxes*; Right: An open *lot box* with wafers

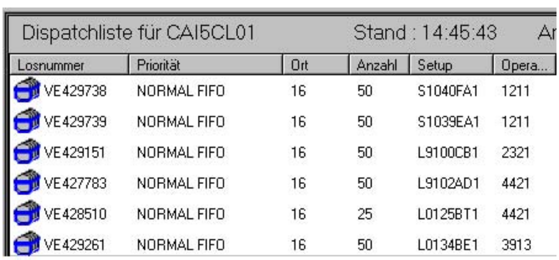
or later (1), and secondly he has to decide to which equipment he delivers the chosen *lot box* (2).

(1) The decision which *lot box* to process next is dependent on the priority of the *lot*. Basically two different types of *lots* may be distinguished. Some *lot boxes* are labelled with high priority such as “express lots” or so called “hand to hand lots”, which have the highest priority. These *lots* have to be handled first by the operator. However, the majority of the *lots* have no designated priority label. Basically the operator is free in his decision in which order to process these *lots*. Nevertheless operators are obliged to follow an implicit ranking based on various information such as holding time (time spent in the shelf) or type of next process step. This ranking is visualized in a so called “dispatch list” (see figure 3), which is explained in more detail later.

(2) The decision on which equipment a specific *lot box* is processed is dependent on the status of the equipment (idle, busy, error, etc.) and the next required program step of the *lot*. The fact that a certain *lot box* can be processed at different equipment for the same program step and the various different priorities make this a non trivial task for the operator.

Here we want to emphasize the importance of a certain freedom of the operator in the choice of the next *lot* and equipment due to flexibility reasons (e.g., breakdown of equipment). To support the operators in their tasks and decision-making process various interfaces can be distinguished. In the following we describe the interfaces the operators work most with, namely *equipment specific interfaces* and *signal lights*, the *FabCockpit* and the *dispatch list*, and the *DisTag*.

Every piece of equipment has specific interfaces and signal lights attached to it. Equipment specific interfaces are numerous and different for every kind of equipment as they range from touch-screens, status LEDs, various buttons to acoustic signals and monochrome (old-fashioned) screens. Signal lights are situated above each equipment (see figure 2) and provide a view over the equipment status. A green light for instance symbolizes that the equipment is processing whereas a red light symbolizes an error. A constant yellow light signals an idle equipment and the completion of processing is shown by a flashing yellow light. Additionally to specific interfaces and signal lights every equipment has a display showing the so-called "FabCockpit". This program was developed to provide a standardized interface in addition to the various heterogeneous types of equipment specific interfaces. The *FabCockpit* provides information about the equipment status as well as a ranked list of *lot boxes*, which can be processed on this particular equipment called the *dispatch list* (see figure 3). The list is generated dynamically and the uppermost *lot box* displayed should be processed first on this equipment. It is possible that one and the same *lot* may be displayed on various lists since this *lot* can be processed on all of these equipment. Therefore, the ultimate decision which *lot* to choose first is still in the hands of the operator. The "DisTag" is a small screen attached to each *lot box*. It provides information about the next process step required, the *lot* number as well as other data regarding this *lot*. This information is used by the operator to distribute *lot boxes* to the various equipment.


Dispatchliste für CAI5CL01 Stand : 14:45:43 Ar

Losnummer	Priorität	Ort	Anzahl	Setup	Opera...
VE429738	NORMAL FIFO	16	50	S1040FA1	1211
VE429739	NORMAL FIFO	16	50	S1039EA1	1211
VE429151	NORMAL FIFO	16	50	L9100CB1	2321
VE427783	NORMAL FIFO	16	50	L9102AD1	4421
VE428510	NORMAL FIFO	16	25	L0125BT1	4421
VE429261	NORMAL FIFO	16	50	L0134BE1	3913

Fig. 3. The *dispatch list* showing *lot* numbers in ranked order

The above mentioned interfaces support the operators in their tasks and decision-making process. Nevertheless our studies showed some potential to improve the work flow towards the "zero-defect" and maximal throughput paradigm. In the following we describe these shortcomings as a motivation for the design of a novel interface informing and guiding the operator, we call *Operator Guide* and provide an initial design.

5 Operator Guide Design

As described above operators have to make the decision which *lot* to process next and on which equipment. These decisions are influenced by two main factors. On the one hand, the operator has to achieve a high throughput which means to process as many *lots* as possible and therefore reduce equipment idle times. On the other hand operators are obliged to follow the ranking provided by the *dispatch list*, in the following referred to as "compliance". A high compliance means a strict adherence of the *dispatch list*, whereas a low compliance indicates that the operator does not adhere the *dispatch list*. These two decision factors may contradict each other since a high compliance may lead to longer equipment idle times as the operator might be forced to wait for high priority *lots*.

One of our major findings was the lack of compliance as the operators maximized their individual throughput because this was assumed as their main evaluation criteria. Nevertheless, from a more holistic perspective, the sum of these local maxima does not necessarily lead to a global optimum when looking at the factory as a whole. An locally optimized equipment load may lead to a bottleneck at some of the further processing steps, causing problems for the entire factory along the remaining path of the *lot box*. Due to the high complexity of the factory these problems cannot be anticipated by individual operators due to a lack of appropriate information.

Another finding was the importance of contextually appropriate information at the right moment. In our study we observed that operators made the decision which *lot* to take next at the delivery shelf for incoming *lots*. At this site no *dispatch list* was available and therefore the operator had no ranking information available. Additionally the operator had no information available on which equipment the next *lot* should be processed. Last but not least we noticed that operators worked with numerous interfaces that caused an information overload. For instance the *DisTag* and *FabCockpit* not only include for this moment relevant information, but provide a lot of information that is not needed by the operators for most of their tasks. Therefore one main design criteria for the *Operator Guide* was to keep it simple by visualizing only relevant information. In sum the goals for the design goals for the *Operator Guide* are

1. a higher compliance by guiding the operator through the optimal next working steps
2. an improved efficiency by simplifying the interface and therefore reducing the information overload, and
3. the location of the interface on each place the operator has to make a decision.

Figure 4 shows a design prototype of the *Operator Guide*. In order to improve learnability we took advantage of mental models of two already deployed interfaces. On the one hand we based the design on the colour scheme of the signal lights described in chapter 4. Therefore red symbols signal an error, yellow icons imply that the operator has to take action and a green colour visualizes a processing equipment. On the other hand the ranking order of the *dispatch list* is utilized. Analogue to it the uppermost line shows the next tasks to be done by the operator.

Each line in the *Operator Guide* represents one task and consists of a symbol that indicates the type of the task, the ID of the equipment on which the task should be performed and (optionally) a *lot* number. In the following we will explain the meaning of each line in the *Operator Guide* shown in figure 4.

1. The uppermost line shows that there has been an error at equipment "F02:136". This is indicated by the red colour and the symbol "X". The operator is obliged to solve the error or to inform the maintenance.

Fig. 4. Design prototype for the *Operator Guide*

2. The next line shows a symbol with an up and down arrow, an equipment ID as well as a *lot* number. This line signals, that the specified equipment has finished processing and is idle (yellow colour). The symbol on the left side indicates that the finished *lot* can be unloaded and a new one loaded. The *lot* number ("RU 751293") in the blue field on the right side visualizes the optimal next *lot* to be loaded into the specified equipment ("F02:139"). The *lot* number matches the ID visualized on the DisTag.
3. The third line is similar to second one except the symbol on the left side. The down arrow means that the equipment is ready to be loaded. Equipment ID ("F02:263") and *lot* number ("RU 915518") have the same meaning as above.
4. The fourth line shows that a *lot* can be unloaded (indicated by the up arrow) at the specified equipment ("F04:131"). Since no *lot* is ready to be processed on equipment no *lot* number is shown.
5. The right arrow symbol in the fifth line indicates that the operator can carry the specified *lot* ("RU 801250") from the delivery shelf for incoming *lots* to the specified equipment ("F04:133"). There the *lot* has to wait until the equipment is idle.
6. The circle arrow and the green colour in the sixth line symbolizes a processing equipment. This means that the operator does not have to take actions at this moment.

The usage of the *Operator Guide* is simple: the operator should perform the uppermost displayed task next. Due to increased flexibility and error prevention not only the very next step, but the next few steps are shown (for instance in figure 4 the next six steps are shown). The list of tasks visualized in the *Operator Guide* is generated dynamically, which means that after completion of each task the corresponding line vanishes. The presented design is the result of three iterations. After each iteration we evaluated the design using a set of focus groups as an expert evaluation method. Both HCI experts as well as experts from the semiconductor factory discussed the design on the basis of usage scenarios.

Currently a prototype of the *Operator Guide* is implemented and will be deployed in one area of the semiconductor factory in the near future.

The *Operator Guide* is an approach to improve the work flow and deal with the shortcomings mentioned in the last chapter. For the design of the *Operator Guide* we followed principles already developed in the field of persuasive technologies. Fogg's [4] conceptional framework, the *functional triad*, distinguishes three roles a computing technology can be: tool, media and social actor. In this sense the *Operator Guide* is designed as a persuasive *tool*. Fogg states that a tool can be persuasive by making the target behavior easier to do or by leading people through a process. The *Operator Guide* increases capability in both ways. It informs the operator to make better decisions by providing a visualization of the next best working steps in a ranked order.

This is achieved by taking three major persuasive principles into account: (1) reduction, (2) tunneling and (3) suggestion [4]. (1) The principle of reduction is to use computing technology to reduce complex behavior to simple tasks, and therefore to increase the benefit/cost ratio of the behavior. The *Operator Guide* follows this rule by minimizing information overload by displaying only relevant information for the next tasks. (2) Another principle is tunneling, which refers to leading users through a predetermined sequence of actions or events, step by step. This is achieved by providing the next best steps along with the tasks to do visualized in a simple way within the *Operator Guide*. Each line of the operator guide provides the relevant information for the working task. It supports compliance as the *Operator Guide* makes it easier to work efficiently by simply following the order provided by the *Operator Guide*. (3) The principle of suggestion signifies that a computing technology will have greater persuasive power if it offers suggestions at opportune moments. The *Operator Guide* provides the right information at the right time, since it is highly contextualized in time and space. It provides contextually appropriate information as it, for instance, shows relevant information at the delivery shelf for incoming *lots* that makes it possible for the operators to choose the *lot* with the highest priority first.

The *Operator Guide* is also a good example for a real world implementation of an ambient display. It is implicitly interwoven with its environment providing only relevant information based on its context. It takes advantage of already built mental modals (e.g. signal light metaphor) and supports the decision making process. At any time only relevant information is shown taking relevant context information and the user into account.

6 Conclusion and Future Work

In this paper we presented the *Operator Guide* as an ambient persuasive interface in the factory context. We have employed a user-centered design approach and described the first two phases in this paper. As a method of study, ethnography has proven to be an appropriate approach for the user requirement analysis phase. Without experiencing the semiconductor factory first hand, including the hands on work in the clean room as part of this particular microcosm, the insights presented in this paper would have been hardly possible. The gained knowledge of the work processes and the used interfaces informed the design of the *Operator Guide*. Combining elements of ambient and persuasive technologies the overall goals of improved work efficiency and reduced error rate have been addressed. The interplay of ambient intelligence paradigms and persuasive strategies results in the *Operator Guide* as a real world example of an ambient persuasive interface. The next steps within the user-centered design cycle will be the deployment of this interface within a specified area of the semiconductor factory. This will allow us to conduct in-situ user studies in order to validate the usefulness, usability and user experience of the *Operator Guide*.

Acknowledgments. The financial support by the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged (Christian Doppler Laboratory for "Contextual Interfaces").

References

1. Aarts, E., Marzano, S.: The New Everyday: Views on Ambient Intelligence. 010 Publishers, Rotterdam (2003)
2. Bertelsen, O.W., Nielsen, C.: Dynamics in wastewater treatment: a framework for understanding formal constructs in complex technical settings. In: ECSCW 1999, pp. 277–290. Kluwer Academic Publishers, Norwell (1999)
3. Fallman, D.: Wear, point, and tilt: designing support for mobile service and maintenance in industrial settings. In: DIS 2002, pp. 293–302. ACM, New York (2002)
4. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. In: Interactive Technologies. Morgan Kaufmann Publishers Inc., San Francisco (2003)
5. Holtzblatt, K., Jones, S.: Contextual Inquiry: A Participatory Technique for System Design. In: Participatory Design: Principles and Practice, pp. 177–209. Erlbaum, Mahwah (2003)
6. Hughes, J.A., Randall, D., Shapiro, D.: Faltering from ethnography to design. In: CSCW 1992, pp. 115–122. ACM, New York (1992)
7. Meschtscherjakov, A., Reitberger, W., Mirlacher, T., Huber, H., Tscheligi, M.: Amiquin - an ambient mannequin for the shopping environment. In: AmI 2009, pp. 206–214. Springer, Heidelberg (2009)
8. Pipek, V., Wulf, V.: Pruning the answer garden: knowledge sharing in maintenance engineering. In: ECSCW 2003, pp. 1–20. Kluwer Academic Publishers, Norwell (2003)
9. Reitberger, W., Tscheligi, M., de Ruyter, B., Markopoulos, P.: Surrounded by ambient persuasion. In: CHI 2008, pp. 3989–3992. ACM, New York (2008)
10. Reul, M.: Bringing usability to industrial control systems. In: CHI EA 2009, pp. 3335–3340. ACM, New York (2009)
11. Yarin, P., Ishii, H.: Touchcounters: designing interactive electronic labels for physical containers. In: CHI 1999, pp. 362–369. ACM, New York (1999)

Reduction of Driver Stress Using AmI Technology while Driving in Motorway Merging Sections

Kashif Zia, Andreas Riener, and Alois Ferscha

Johannes Kepler University Linz, Institute for Pervasive Computing,
Altenberger Str. 69, A-4040 Linz, Austria
Tel.: +43/732/2468–1432; Fax: +43/732/2468–8426
lastname@pervasive.jku.at

Abstract. High average intensity of traffic and problems like traffic congestions, road safety, etc. are challenging problems striking highway operators in these days. With the broad application of intelligent transport systems (ITS), particularly for the most dense street sections, some of these problems can be minimized or even solved; supplementary great potential is attributed to applications based on state-of-the art technology like car-to-x communication, for instance by extending an individuals "field of vision" by observations taken from all the vehicles in front. In this work we present a simulation based approach for improving driving experience and increasing road safety in merging sections by redirecting vehicles in advance according to a negotiation of requirements and desires of the flowing traffic on the main road and cars merging from the entrance lane. The simulation experiments performed in a cellular automaton based environment were data driven and on real scale, using traffic flow data on a minute-by-minute basis from a large urban motorway in a main city of the European Union. Our results have shown that the application of AmI technology has potential to influence driver's behavior (seamlessly invoking for a lane change well before an abrupt merging point) resulting in a reduction of panic, particularly for sections with limited range of view.

Keywords: Data driven simulation, Driver assistance, Motorway merging, Field of view extension, Vibro-tactile seat, Safety belt interface.

1 Introduction and Motivation

It has been reported for merging sections of roads to have a high influence on the overall road safety as different drivers (e. g. that driving on the entrance ramp or on the main roadway) may behave, dependent on state of traffic and their physiological state, in different ways [8] and/or have opposed desires. A driver's purpose entering a main road is to do so without any delays, thus merging into immediately – which could demand traffic on the main road to apply the brake to avoid rear-end collision accidents. On the other hand, traffic moving on the

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 127–137, 2010.

main road also want to keep on moving with desired individual speed and without consideration of cars entering the road and causing changes in their driving habit. This, in turn, may affect fluidity of entering traffic as they would have to wait until a large enough free gap between two cars to merge into. For roads with two or more lanes (which is the focus of this paper), drivers would also have to change lanes in order to (*i*) reach the lane merging into the desired road (merging car) or (*ii*) make space for merging cars to do so (cars on the main road). For the latter group it has been reported, e. g. in [10], that sometimes drivers panic when they want to change a lane leave the motorway or let merging cars into. Buld *et al.* [1] have conducted driving simulator studies at entrance ramps of motorways to investigate the effect of traffic intensity on driver stress. From their results it can be derived that entrance ramps generates already high driver stress for medium volume of traffic and that the effect is increased for unknown ramps. Such situations could be avoided when, by application of ambient intelligence (AmI) technology, seamlessly requesting the driver to change into the appropriate lane early – this means 100s of meters ahead the merging in fluid driven road sections with large gaps. We hypothesize that the application of this approach using a background intelligence available to all cars and broad availability of state-of-the-art wireless communication technology (car-to-car communication) allows for information delivery to a driver from cars hundreds of meters ahead the field of view of the driver. The so increased "visual range" allows for earlier reactions on traffic conditions or infrastructure characteristics., would increase traffic fluidity at merging sections and furthermore decrease driver panic.

The motivation of our approach follows from the fact that many people are afraid about late merging (exceptionally in crowded situations); however, often it is not possible to change lanes earlier (for instance when unfamiliar with a road segment) because of unseeable ramps or constrictions (merging of two lanes into one; blocked lane because of road works or accident). An early realization about road structure, merging lanes and on and off ramps, forcing them to change into the appropriate lane as early as possible would ensure an improvement. A early notification coupled with application of AmI technology would guarantee that the driver is informed inattentive. The technology of choice consists of a tactile car seat, with vibration elements stimulating a driver's skin to guide him/her about lane changes (towards left or right to be particular). According to an observation of real traffic data from a large urban motorway in a city of the EU and with a very high traffic density of about 300,000 vehicles a day we suppose that a early lane change notification could help to reduce overall panic in and around merging sections. Additionally, the seamless skin based ambient vibration system guarantees the delivery of this notification in a non obstructive way.

Outline: The rest of the paper is structured as follows. In section 2 the developed AmI technology for assisting drivers' is explained in detail, section 3 focuses on the simulation model, with emphases on data and simulation environment/methodology. In section 4 the simulation results are discussed, section 5 finally concludes the paper.

2 AmI Technology: A Vibro-Tactile Driver Seat

Both visual and auditory channels of sense are highly saturated in vehicle operation for fulfilling driving related (primary) tasks and side activities; for a successful transmission of additional information, e. g. on lane changing or speed adaptation, it would be required to use (or even abuse) information capacity still unused. In addition it has also to be revealed that the amount of "free space" is (*i*) dynamically changing and (*ii*) dependent from the current traffic situation; thus, there is no guarantee for free capacity at all.

The sense of touch is, with about 10% of the entire information transmission capacity of humans, ranked third behind visual (70–80% [9], [5]) and auditory (10-15% [2]) senses; regardless, tactile interfaces are still used rarely in the automotive domain so that we can infer potential for this sensory channel to deliver the required information accurately and at all times. Our presumptions are supported by many recently presented applications and research studies using vibro-tactile (driver) notification, e. g. [6], [3], [7], [11]. Vibrations can be delivered precise with regard to the position (point-like) and there are many possibilities for varying the stimulation signal (frequency, intensity, pulse-pause interval, combination of vibrations using many elements, vibration patterns or tactograms [11, p. 123]). Additionally, small vibro-tactile transducer could be easily integrated or embedded into controls elements of a vehicle, such as the steering wheel, gearshift, etc. but more universally into parts the driver is in contact all the time, like safety belt or seating. Therefore, we recommend the application of a tactile user notification system integrated into either the driver seat or the safety belt, and subtle notifying the driver on merging related information collected from a background intelligence (such as shown, for instance, in [4]). The mapping of activities with the vibro-tactile notification system is proposed as follows:

(i) **Lane changing: Move/merge to the lane on left**
– *Seating:* All tactors on the left are activated; varying pulse-pause time corresponding to the importance for the driver to perform the action.
– *Safety belt:* Tactors activated one after the other in a "running light" style from bottom to top; varying sequence speed corresponding to the importance for the driver to perform the action (the shorter the distance to the merging point, the faster the running light).

(ii) **Lane changing: Move/merge to the lane on right**
– *Seating:* All tactors on the right are activated; varying pulse-pause time as above.
– *Safety belt:* Tactors activated one after the other in a "running light" style from top to bottom; varying sequence speed as above.

(iii) **Speed variation: Increase driving speed**
– *Seating:* Running light from back to front for both the left and right strip (see Fig. 1); running light speed corresponding to the deviation in speed.
– *Safety belt:* "Outer elements" (e. g. tactors 1, 2, and 7, 8) are activated until the expected target speed is reached; vibration intensity corresponding to the speed deviation.

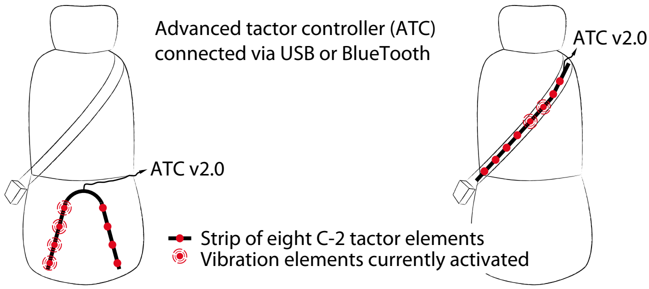


Fig. 1. Vibro-tactile stimulation in the seating, the safety belt or other control instruments in the car is projected to be used for (inattentive) notifying the driver on upcoming merging-related operations

(iv) **Speed variation: Decrease driving speed**
- *Seating:* Running light from front to back for both the left and right strip; rest as above.
- *Safety belt:* 'Inner elements" (e. g. tactors 3 to 6) are activated until the expected target speed or zero is reached; vibration intensity corresponding to the speed deviation (as above).

3 The Simulation Model

Data and Environment: The agent-based modeling method is applied as simulation paradigm using NetLogo [13] as a tool. NetLogo is a cellular automaton (CA) based simulation tool in which a grid of placeholder agents (patches) define an environment (world). The movable agents (turtles) reside on top of the patches. Each of the patch is defined by a x and y coordinate as that of a turtle residing on it (if any). For the purpose of traffic simulation, a single patch is assumed to have a width equal to $5m$. Given that a single vehicle can occupy a patch at a time, this ensures an average vehicle length of 5 meters considered to be a constant, applicable throughout the simulation. Modeling of different types of long vehicles like buses, trucks would be possible by a combination of two or more turtles, hence obtaining length of $10m$, $15m$, etc. On the other hand, the height of the patch represents the width of the lane a vehicle is moving on (this value has been devoted scant attention as it has almost no influence on the fluidity of traffic or forward moves).

The processed simulation was running data driven (with regard to the traffic flow), with an underlying road network model built true to scale from a segment of a large urban motorway in a main city of the EU (the average intensity of traffic on this road is higher than $300,000$ vehicles a day). Due to the size of the network, with nodes of roads may having up to sixteen lanes distributed in different levels, it has become one of the most traveled routes within the EU and a (inter)national reference as object for traffic studies. The focused segment contains the "main road" absorbing a merging lane, an offshoot from the "lower road". The data collection agency collects the induction loop sensors data at different points of measurement (PM[1]). This data is minute-to-minute based focusing on vehicle count and average speed along with other measures. The induction loop sensors related with the road segments are shown in Fig. 2.

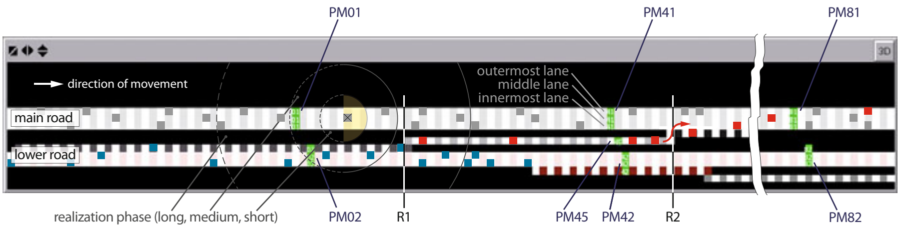


Fig. 2. "Realization" corresponds to an extended field of vision for a driver, obtained with help of AmI technology. PMxx represents the traffic counters of the road segment.

The following subset of the measurements are considered for the simulation:

- Vehicle count at PM01: Vehicles/hour counted at PM01 each minute, necessarily to generate traffic flow on "main road".
- Vehicle count at PM02: Vehicles/hour counted at PM02 each minute, necessarily to generate traffic flow on "lower road".
- Vehicle count at PM45: Vehicles/hour counted at PM45 each minute, necessarily to generate traffic flow on "merging lane"; this essentially should be a subset of vehicle count at PM02.

The purpose of the study is to ensure a smooth and less panicked merging of traffic from the merging lane to the main road following two considerations:

- For the smooth merging of the merging lanes vehicles into the main road, the innermost lane of the main road is most important. If there is a gap between incoming traffic on this lane, it would help smoothen the merging. But a totally vacant lane can also produce a lot of load on the two upper lanes, thus destroying the overall effect.
- The vehicles coming from PM02 should occupy an appropriate lane before reaching to R1 and vice versa. If drivers perform preemptive lane changes this would ensure less panic near R1 (due to the absence of congestions).

The one factor directly influencing the above two considerations is, when a driver on the main road realizes that he/she has to perform an action due to a incoming merging ramp. For example, the following actions are imaginable:

- Drivers on the main road move up in their lanes to facilitate merging traffic.
- Drivers on the main road which could not move up in the lanes (e. g. due to heavy traffic) and confront a merging car on the ramp extension may accommodate that vehicle in opposition to their right.

[1] PM . . . measurement point ([french] point de mesure).

– Drivers on the lower road intending to merge move up in the lanes to facilitate themselves.
– Drivers on the lower road intending to keep on moving on this road move down in the lanes to facilitate themselves.
– In case of interest conflict near R1, there is no logical game theory (strictly local) which can safeguard interests of both the confronting parties.

To accommodate all of the above actions, an underpinning is the time when a driver realizes existence of a ramp at a distance. Throughout this study, we have focused on this factor named as "realization range". With a variation in realization (assisted by AmI assistance), we have compared the average speed and appropriate lane positioning to figure out the best suitable combination.

Simulation Setup: The overall simulation process is represented with the flow shown in Fig. 3. From the dataset relating to the PMs in focus (as shown in Fig. 2), we have taken data rows representing one day of activity. Considering that the data was provided on a minute-to-minute basis this equals to 1,440 records. In the simulation flow, we execute (read) data contents from the truncated file until the end of file marker is reached. Data preprocessing works as follows. Each iteration of the simulation represents a single second. For that it is required to firstly convert the vehicle count given on a vehicles/hour scale to a vehicles per minute scale, and then to randomize the traffic for a single minute for each of the simulation second. So, in reality, in each second (iteration step) of a minute (data row), zero to three vehicles would be registered in datasets relating to PM01 and PM02 readings. After having second to second traffic generation datasets, the simulation is run for a whole minute (60 iterations). In each second, the vehicles are generated and moved forward. In case of the first 60 seconds of the "life of a vehicle" it would be considered as a new vehicle and would not experience any lane change or speed variation activity. In fact, this information is embedded into a more complex "move forward" process flow as illustrated in Fig. 4.

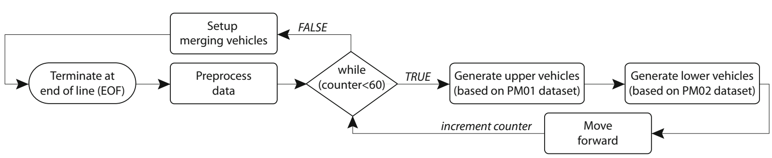


Fig. 3. Overall simulation flow (main process)

Assess vehicle states: A relatively simple task is to categorize the vehicles present in the simulation after simulation of (the first) 60 seconds. The corresponding process flow is represented by the process "Setup merging vehicles" – which is all about managing different vehicle states (or colors) as elaborately discussed subsequently. The default color of vehicles at the lower road is blue whereas the default color on the main road is gray. Both blue and gray agents represents vehicles keep on moving on their roads (without merging; note that there are no merging vehicles at the main road). Internal to the process, the flag "new-vehicle" represents a new vehicle which is not set for merging. For this reason all the vehicles in the lower road generated during the latest minute are marked "new", and as soon as the first 60 seconds are over, a comparison of all the new vehicles with the corresponding amount of merging vehicles at PM45 is made (PM values obtained from the data table). If the number of vehicles merging in the real (PM45 reading) is equal or greater then the number of new vehicles generated at PM01, all new vehicles have to be set as having the intent to merge, thus setting their color to red (a PM45 reading higher than the number of generated vehicles may appear due to data error.) If, in contrast, the number of new vehicles generated (PM01) is greater then merging vehicles (PM45), then exactly that number of vehicles equal to the PM45 values is set to merging vehicles (=red), the remaining are set to "keep on moving" (blue). At the end of this process flow, all newly generated vehicles have obtained a status/color (either red or blue) according to fractions from real vehicle movement behavior.

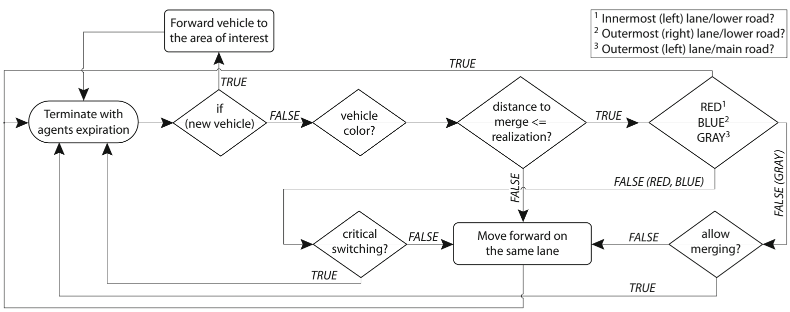


Fig. 4. Process of vehicle movement during a single second of simulation

Vehicle movement: The set of processes representing the move forward activity (see Fig. 4) is applied to all the agents (vehicles) in the system in each iteration (second). If the agent in consideration is a new vehicle, we just forward further until it reaches the area of interest (see Fig. 2). When it reaches there (influenced by the range of realization actually used for simulation), it is no longer considered as a new vehicle (applies to both lower and main roads); when moving on the lower road, additionally its merging nature (color) is set. For a vehicle which is not new, there are three possible vehicle states represented by vehicle colors, (*i*) merging (red), (*ii*) keep on moving on the lower road (blue), and keep on moving on the main road (gray). Whatever the state (color) of the vehicle is, first we need to establish its distance to the merging point (R1 for lower road traffic, R2 for main road traffic). As soon as a driver realizes that there is a merging situation ahead, he/she is within merging distance. The earlier a driver knows about an upcoming merging region, the "better" he can react on this situation based on his own desires and observations from the environment. We have experimented with different AmI technology extended "fields of vision" – this distance is represented in the simulation model by the variable *realization*

1) Lower road, vehicles intending to merge into main road (colored red)
A. Red vehicle within merging region and before R1: In this case, the vehicle would try to upgrade its lane. If successful, its execution for this iteration is complete. If it is not able to perform a up gradation (due to insufficient lane change distance or if it is already in the upper most lane), it checks its position. A „critical switching“ situation occurs if the red vehicle is in the middle or lower lane and it is very close to R1. It needs to switch its lane immediately. If critical switch- ing is not required, the vehicle moves forward in the same lane ensuring a incremental speed variation (if required).
B. Red vehicle after R1 but before R3: In this case, if a critical switching situation occurs due to end of ramps R2 or R3, the vehicle needs to switch its lane immediately (moving up). If critical switching is not required, the vehicle moves forward in the same lane ensuring a incremental speed variation (if required).
C. Red vehicle after R3: In this case, there is no need for critical switching and the vehicle moves forward in the same lane ensuring a incremental speed variation (if required).
2) Lower road, vehicles intending to keep on moving on the lower road (blue cars)
A. Blue vehicle within merging region and before R1: In this case, the vehicle would try to downgrade its lane. If it is success-ful, its execution for this iteration is complete. If it is not able to perform a down gradation (due to insufficient lane change distance or if it is already in lower most lane), it checks it position. A critical switching situation occurs, if the blue vehicle is in the upper lane and it is very close to R1. So it needs to switch its lane immediately. If critical switching is not required, the vehicle moves forward in the same lane, ensuring a incremental speed variation (if required).
B. Blue vehicle after R1: In this case, there is no need for critical switching and the vehicle moves forward in the same lane ensuring a incremental speed variation (if required).
3) Main road, vehicles intending to keep on moving on the upper road (gray cars)
A. Gray vehicle within merging region and before R2: In this case, the vehicle would try to upgrade its lane. If it is successful, its execution for this iteration is complete. If it is not able to perform an up gradation (due to insufficient lane change distance or it is already in upper most lane), the vehicle moves forward in the same lane, ensuring a incremental speed variation (if required).
B. Gray vehicle within merging region and within R2 and R3: In this case, the vehicle would try to upgrade its lane. If it is successful, its execution for this iteration is complete. If it is not able to perform an up gradation (due to insufficient lane change distance or if it is already in upper most lane), there are two further possibilities. First, when it is very close to R3 and is in the lower lane it will check its vicinity. If a red vehicle is waiting in the merging lane just before R3 it allows the vehicle to merge by slowing down (or stopping) itself. If a „allow merging situation“ is not required the vehicle moves forward in the same lane ensuring a incremental speed variation (if required). The second option is, when it is not very close to R3, that it moves forward in the same lane ensuring an incremental speed variation (if required).
C. Gray vehicle after R3: In this case, the vehicle moves forward in the same lane ensuring a incremental speed variation.

Fig. 5. Merging possibilities according to vehicles' distance to the point of merging

(explained in Fig. 2). For simulation, one important factor to cope with is the difference between a merging point already passed and a merging point located ahead. Relating to Fig. 2, in Fig. 5 all of the implemented possibilities are listed. It is important to note that successful lane changing would require a appropriate lane change distance. A second factor is that when a vehicle is moving in the same lane it performs an incremental or decremental speed variation all the time. If its speed is slower then the maximum speed allowed it would increment its speed up to the maximum but with ensuring a safe breaking distance from the vehicle in front. If it needs to apply breaks it would decrement its speed. These functionalities are, of course, part of the simulation and explained elsewhere [12].

4 Simulation Results and Discussion

The measurement points PM01, PM02 and PM81 were considered for measuring the traffic flow – it can be observed that the average speed of traffic is inversely proportional to the realization range (Fig. 6). Furthermore, the bottleneck at PM81 does not allow much variation in speed for a long simulation

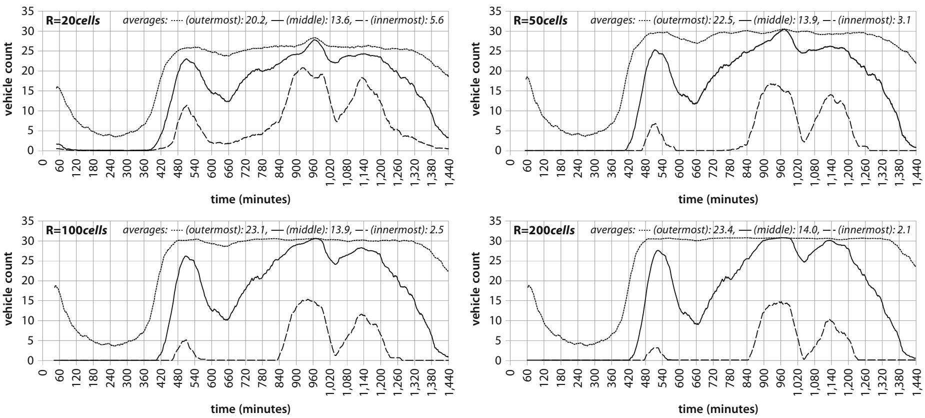


Fig. 6. Lane positioning at PM41/main road (only "gray" vehicles are counted)

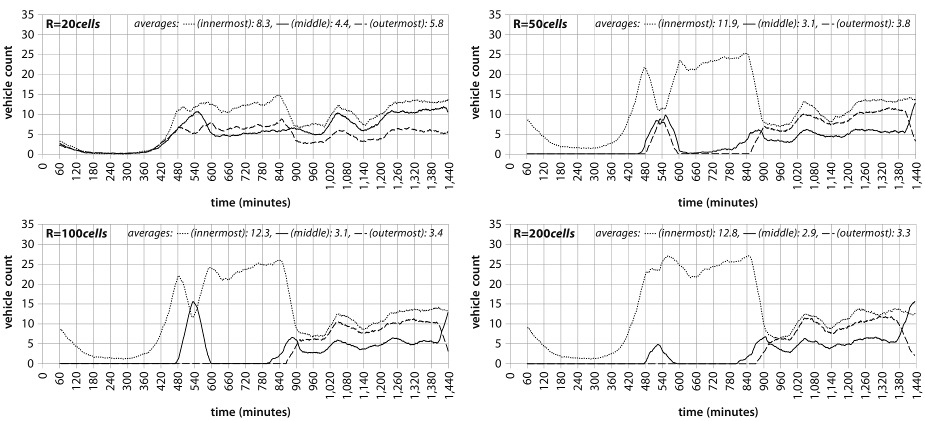


Fig. 7. Lane positioning at PM02/lower road ("red" vehicles only)

and is seen to influence the flow between PM01/PM02 and PM81 transferring the congestion in the backward direction. Having established that the overall throughput of the road section (governed by output point PM81 after merging) would not be diversely affected by realization range, we proceed to ask the question addressed by this paper, i.e. would panic be reduced with an increase in realization range? The graphs shown in Fig. 6 and Fig. 7 respond affirmatively. PM41 and PM02 are used to observe the lane occupations of vehicles where all traffic is notified about the merging just ahead. In case of main road (PM41), the innermost lane ideally should be vacant just before R2 to allow unrestricted merging. For realization range equal to $100m$, the innermost lane accommodates $5.6 vehicles/minute (v/h)$, which is progressively reduced to $2.1 v/m$ with an increase in realization range. The middle lane accommodates almost the same number of vehicles for all four cases, whereas the load from innermost lane shifts directly to outermost lane. Similarly, at PM02 the red vehicles should occupy the innermost lane just before merging. Due to too less time to perform lane changes in case of realization range equal to $100m$, less then half the red vehicles are residing in innermost lane (8.3 out of $18.5v/m$. However, with increasing the realization range this measure progressively increases to $12.8v/m$, influencing a reduction in the load of outermost and middle lanes [1]. As suggested by Buld [1], the presence of an abrupt and unknown ramp increases the driver's stress (panic). An early notification in an un-obstructive and seamless manner under the influence of an ambient setting would reduce the panic given that the driver follows the suggestion which should result in moving most of the vehicles in optimal lanes well before merging. The simulation results show the potential of panic reduction even in a high traffic road section.

5 Conclusions

Considering vibration based ambient technology providing necessary information relating to the road structure (global) and vehicles neighborhood (local) to a driver seamlessly, a data driven simulation was performed to analyze the effectiveness of two kind of notifications relating to lane and speed changes. While the speed change notifications do not generate the similar effects as that of real data values, the proposed lane changing strategy guarantees a less panicking setting near and on the merging points. As the realization range (a look ahead distance a driver is notified about an incoming merging situation) increases, the forced vibration based notification systems forces the driver to merge into appropriate lane. This would ensure an ease in highly congested merging situation, consequently decreasing the potential panic if a substantial amount of drivers find themselves into wrong lanes just before merging.

Acknowledgements. This work is supported under the FP7 ICT Future Enabling Technologies program of the European Commission under grant agreement No 231288 (SOCIONICAL). We would further like to acknowledge the provider of traffic data (who prefers to remain anonymous).

References

1. Buld, S., Hoffmann, S., Totzke, I., et al.: Adaptive driver assistance on the basis of traffic condition – exemplary demonstration for highway entrance. In: Aachener Kolloquium Fahrzeug- und Motorentechnik, ch. 15, pp. 1758–1782 (2006)
2. Dahm, M.: Grundlagen der Mensch-Computer-Interaktion, 1st edn., 368 pages. Pearson Education, London (2005)
3. Erp, J.B.V., Veen, H.A.V.: Vibrotactile in-vehicle navigation system. Transportation Research Part F: Traffic Psychology and Behaviour 7(4-5), 247–256 (2004)
4. Ferscha, A., Riener, A.: Pervasive Adaptation in Car Crowds. In: First International Workshop on User-Centric Pervasive Adaptation (UCPA) at MOBILWARE 2009, Berlin, Germany, April 27, p. 6. Springer, Heidelberg (2009)
5. Hills, B.L.: Vision, mobility, and perception in driving. Perception 9, 183–216 (1980)

6. Ho, C., Tan, H., Spence, C.: Using spatial vibrotactile cues to direct visual attention in driving scenes. Transportation Research Part F: Psychology and Behaviour 8(6), 397–412 (2005)
7. Kwon, D., Kim, S.: Haptic Interfaces for Mobile Devices: A Survey of the State of the Art. Recent Patents on Computer Science 1(2), 84–92 (2008)
8. Liu, R., Hyman, G.: Towards a generic guidance for modelling motorway traffic merge. In: European Transport Conference (ETC), Leiden, Netherlands, October 6-8, p. 17. Association for European Transport, AET (2008)
9. Mauter, G., Katzki, S.: The Application of Operational Haptics in Automotive Engineering. In: Business Briefing: Global Automotive Manufacturing & Technology 2003, pp. 78–80. Team for Operational Haptics, Audi AG (2003)
10. New Jersey Motor Vehicle Commission: New Jersey Driver Manual, ch. 5: Defense Driving (2008)
11. Riener, A.: Sensor-Actuator Supported Implicit Interaction in Driver Assistance Systems, 1st edn. Vieweg+Teubner Research, Wiesbaden (January 14, 2010) ISBN-13: 978-3-8348-0963-6
12. Riener, A., Zia, K., Ferscha, A.: AmI technology helps to sustain speed while merging – A data driven simulation study on Madrid motorway ring M30. In: DS-RT 2010, Fairfax, VA, USA, October 17-20, p. 10. IEEE CS Press, Los Alamitos (2010)
13. Wilensky, U.: Netlogo modeling environment, http://ccl.northwestern.edu/netlogo, (last retrieved July 30, 2010)

Subjective Difficulty Estimation for Interactive Learning by Sensing Vibration Sound on Desk Panel

Nana Hamaguchi, Keiko Yamamoto, Daisuke Iwai, and Kosuke Sato

Graduate School of Engineering Science, Osaka University,
Machikaneyama 1-3, Toyonaka, 5608531 Osaka, Japan
hama-n@sens.sys.es.osaka-u.ac.jp, kei@kit.ac.jp,
{daisuke.iwai,sato}@sys.es.osaka-u.ac.jp

Abstract. In this paper, we propose a method which estimates the student's subjective difficulty with a vibration sound on a desk obtained by a microphone on the back of the desk panel. First, it classifies the student's behavior into writing and non-writing by analyzing the obtained sound data. Next, the subjective difficulty is estimated based on an assumption that the duration of non-writing behavior becomes long if the student feels difficult because he (or she) would not have progress on answer sheet. As a result, the accuracy of the proposed so simple behavior classification reaches around 80%, and that of the subjective difficulty estimation is 60%.

Keywords: subjective difficulty estimation, behavior classification.

1 Introduction

When a student studies alone and works through a lot of topics in a classroom, it is difficult for the student to correctly comprehend the proficiency level of each studied topic by herself (Fig. 1). As an example, let's consider a situation where the student works through dozens of math problems and reaches the correct answers. She will come to the conclusion that she has mastered the material and further review is not necessary. However, what if she can easily answer most of the questions, but has difficulty with others?In such case, she should review the topics of the difficult problems. However, it can be easily imagined that the student is unable to remember the difficulties of each problem herself. It is also difficult for a teacher to know how each problem is difficult for each student because the difficulty is psychological value, thus it cannot simply be measured by the score of an examination. If a teacher spends a substantial amount of time supervising the student, it might be possible to comprehend the subjective difficulty of each problem. However, teachers have to take care of dozens of students, thus it seems hard to do the same thing for all of them.

We give a solution to the issue with our proposed system which can estimate the subjective difficulty with each problem by observing the student's behavior

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 138–147, 2010.

based on the concept of "Ambient Sensing". Our proposed system measures the vibration sound on a desk panel, there by estimating the subjective difficulty of each problem in accordance with the measured sound, then saves it in a PC. Students can browse the subjective difficulty of each problem on a PC, and thus easily notice how well they actually understood the problems. Their teachers can also refer to the visualized subjective difficulty data on the PC.

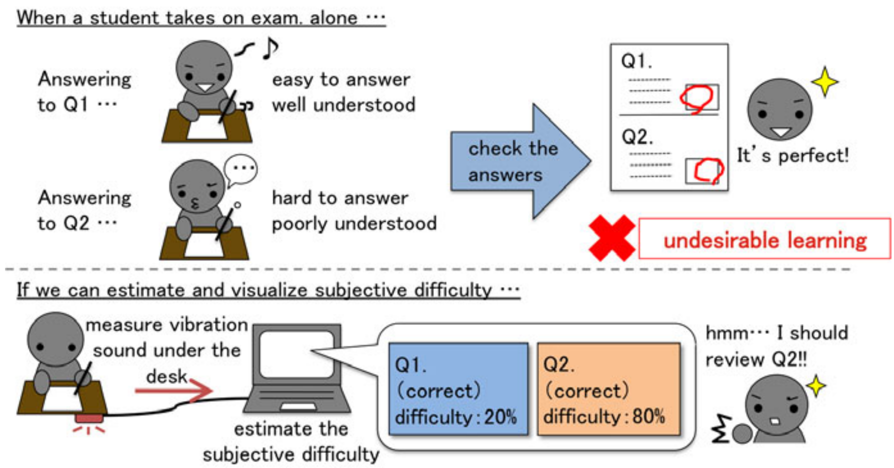


Fig. 1. The problem of studying by oneself and the advantage of our proposed system

2 Related Work

Our research is closely related to works on emotion estimation. Researchers have been investigating various emotions through physiological or behavioral measurements. The physiological measurement is reliable in the emotion estimation. However, this measurement cannot be performed in small classrooms or homes because large equipment is required. Furthermore, because sensors must be attached on the student's body, the sensors disrupt natural study. In the behavior measurement, image sequences captured by a video camera directed toward a student have analyzed the emotion estimation. Kappor *et al.* have investigated the estimated psychological stress of a student[1]. Their proposed system analyzed a student's facial expression, head velocity, and mouth movement from the captured images. Gunes *et al.* used two cameras to classify the participant's emotions into anger, disgust, fear, happiness, uncertainly and anxiety[2]. One was used for capturing a participant's facial expression, and the other was for upper body gestures. With the bi-modal information, they achieved over 94% classification accuracy. As described above, the camera-based approach could work well to estimate various emotions. However, especially in the case of estimating student's subjective difficulty, the approach would have a problem that the presence of camera might disturb the student's studying. Some of student might feel that they are being monitored by someone and consequently feel extra psychological stress.

Other behavioral measurement approach for emotion estimation exist. Okubo *et al.* proposed to estimate how a student concentrates on learning by analyzing data from a 3-axis accelerometer attached on the upper part of the backrest of a chair[3]. They assumed that little movement meant the student was concentrating. Because the student was unaware of the presence of the sensor, he (or she) did not feel any extra psychological stresses. However the system observed only the chair movement, and the obtained data was not sufficient to estimate the subjective difficulty of a question.

Lv *et al.* proposed to apply a pressure sensor keyboard which could measure finger pressures on the keys[4]. They classified the participant's emotion into neutral, anger, fear, happiness, sadness, and surprise from the measured pressure patterns. Luria *et al.* found that there was a difference between the handwriting of true and false sentences by analyzing pen pressure and trajectory measured by a pen tablet [5]. Although they worked well, keyboard and pen tablet are not widely used by students in studying at the moment. Instead, the conventional pen-and-paper based learning is still commonly applied in the school education.

Harrison *et al.* invented the system which was referred to Scratch Input. Scratch Input could detect the sound of scratching a textured surface with a fingernail and recognize several gesture patterns [6]. The authors applied two microphones and obtained the vibration sound occurring on the desk, and then, they recognized the user's gestures by analyzing the obtained sound. This would be able to apply to the study situation where the student uses a pen and paper because the sound of scratching with both a fingernail and a pen seems to be similar. Scratch Input was useful as the behavior estimation but it did not consider about estimating the user's emotion.

Compared to the previous works, we propose to estimate a student's emotion, in particular the subjective difficulty of a studied topic, in such a way that the student neither wears any sensors nor feels to be monitored. In addition, to allow the student to use the familiar pen and paper, the proposed system employs to measure the vibration sound on the desk.

3 Subjective Difficulty Estimation Method

3.1 Human Behavior Distinctive in Working on Difficult Topic

This section describes the proposed subjective difficulty estimation method. At first, we investigate what kind of behavior a student behaves when working on a difficult problem. Previously, Nakamura *et al.* carried out a user study in which a participant was observed when working on some difficult problems with a e-learning system. As a result, the authors found that the participant gazed one point on a PC monitor and did not click frequently for a long time. This result indicates that people tend not to move their bodies when they try to answer difficult problems to increase concentration.

In addition, we conducted a questionnaire survey where five high school teachers participated. On average the teacher have 14 years of experiment. We asked them the following two questions:

Q1. Can you comprehend how well your students understand each topic, only by observing their behaviors in an examination?
Q2. If Q1 is "yes", then which kind of behavior is the most important key for you to estimate their proficiency levels on the topic?

In terms of Q1, four participants answered that they could comprehend their students' proficiency levels, not completely but partially. In terms of Q2, all four teachers told that they focused on how fast their students solve each problem. When students understand a topic, they can answer related problems quickly, and vice versa.

3.2 Ambient Sensing

We apply the "Ambient Sensing" approach to our subjective difficulty estimation system. The "Ambient Sensing" approach measures human behavior by using sensors embedded in a natural environment so that the user does not notice the presence of the sensors. Thus a student neither feels to be monitored nor wears/holds any special devices. The proposed system estimates the subjective difficulty in a usual learning environment where a student uses a pen and paper on a desk.

We focus on the vibration sound on the desk. When a student is writing an answer to a question with a pen, the pen hits the desk through the paper and consequently the desk vibrates. We measure the vibration as sound by attaching a small microphone on the backside of the desk. We analyze the sound data to detect whether the student is writing, and then estimate the subjective difficulty of the current question.

3.3 System Configuration

Figure 2 shows our first prototype system. We prepared a ballpoint pen and a desk (500 $\times$ 400 $\times$ 670 mm) whose top was chiefly made of wood. As described in 3.2, we attached a small piezoelectric microphone (27 mm radius, less than 1 mm thick) on the backside of the desk. A PC analyzed the measured sound data and estimated the subjective difficulty. The sound data was amplified before sent to the PC.

3.4 Handwriting Estimation

We set the assumption that a student repeats two states, the THINKING-STATE and the WRITING-STATE, while studying, as shown in Fig. 3. As a preliminary stage of the subjective difficulty estimation, we propose to distinguish THINKING-STATE from WRITING-STATE by analyzing the measured

Fig. 2. Overview of prototype system

sound data. Output of this processing, "Handwriting Estimation", is handled to the next stage, "Measure Subjective Difficulty", which is indicated in the next section.

Figure 4 shows the classification tree applied to the measured vibration sound data. The measured data is at first classified into the SILENCE class and the SOUND class. The data falling into the SOUND class is then classified as the WRITING-SOUND class and the NOISE class. We only consider the WRITING-SOUND case when collision and friction sounds are made by a pen on paper. The other sound which occurs when the student does not write anything is regarded as the NOISE (e.g., the friction sound of the paper or the student's hand with the desk). We refer to the SILENCE and the NOISE classes together as the NON-WRITING-SOUND class. The proposed sound classification is related to the two states of Fig. 3. The THINKING-STATE corresponds to the NON-WRITING-SOUND class, while the WRITING-STATE corresponds to the WRITING-SOUND class.

The vibration sound on the desk is measured as a monaural signal with the sampling frequency of 44,100 Hz. The measured sound is downsampled to 11,025 Hz, and then its Fourier transform is computed. We analyze the frequency characteristics of the measured sound data to classify it into the proposed sound classes. As a preliminary study, we measured vibration sounds on the desk in the following conditions: (1) a participant wrote texts with a pen and paper on the desk (i.e., WRITING-SOUND), (2) she did not write anything but did other things such as tapping with her finger or sliding the paper on the desk (i.e., NOISE), and (3) nothing existed on the desk (i.e., SILENCE). As a result, we found that there were differences in the spectra at frequency band of 200–250 Hz and 1,050–1,130 Hz between WRITING-SOUND and NON-WRITING-SOUND. Figure 5 shows the spectra of each sound class. Thus, we use the means and variances of the spectra at these two frequency bands to classify the vibration sound

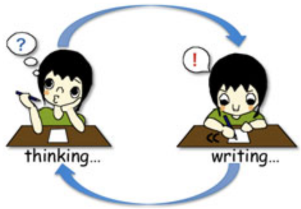


Fig. 3. Student's behavior model when she is working through a problem

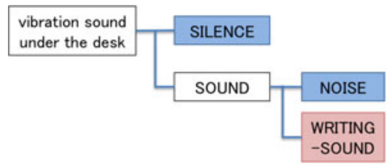


Fig. 4. Classification tree of vibration sound on the desk. We refer to the classes SILENCE and NOISE as NON-WRITING-SOUND respectively, whereas WRITING-SOUND.

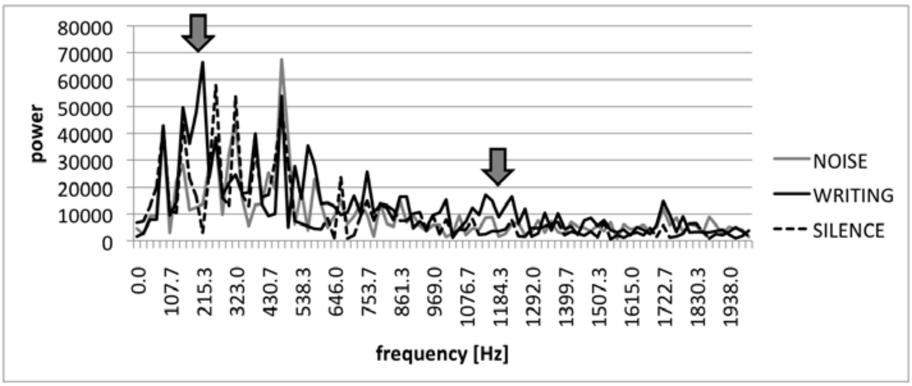


Fig. 5. The spectra of NOISE(gray solid line), WRITING-SOUND(black solid line), and SILENCE(black dot line)

into WRITING-SOUND and NON-WRITING-SOUND. We apply a decision tree for the classification.

3.5 Measure Subjective Difficulty

To investigate the relationship between subjective difficulty and writing behavior, we observed a few studying students working on a series of calculations in a classroom. As a result, they often stopped their writing hands when they worked through difficult problems. This was regarded that the students tend to be more in the THINKING-STATE than in the WRITING-STATE then. Thus, we define the ratio of the duration of the NON-WRITING-SOUND to the total amount of time spent for a question as the subjective difficulty of the question. Suppose that the WRITING-SOUND duration is W_t and the non-writing-sound duration is N_t, the ratio R is computed as:

$$R = \frac{N_t}{W_t + N_t}. \tag{1}$$

4 Experiment

4.1 Evaluation of Hand-Writing Estimation

We constructed a decision tree to classify the captured vibration sound into WRITING-SOUND and NON-WRITING-SOUND based on algorithm C4.5. Five subjects were recruited from the local university (age 24.2 ± 2.2 years, mean ± SD; 3 female, 2 male) and were requested to work on simple calculation problems on a paper using MVPen[1] to answer the problems. MVPen is a ballpoint pen equipped with a sensor which can detect whether it touches the paper, its sampling rate is 211.06 Hz. The touch information obtained from MVPen could be regarded as reference. We prepared sixty different problems for each subjects.

When each subject worked through the calculations, the vibration sounds on the desk were captured by the microphone and the touch information obtained from MVPen were recorded. Time series of the captured sound data was segmented in such a way that the duration of each segment was 1/211.06 s, and then each segmented was classified into the WRITING-SOUND and the NON-WRITING-SOUND classes by using the reference. For each segment, four feature values were calculated: the mean and variance values of the spectra at the frequency bands of 200–250 Hz and of 1,050–1,130 Hz. Using these parameters, we made a decision tree based on C4.5.

We carried out an evaluation of the decision tree. We computed the accuracy of the classification by using the same sound data. Figure 6 shows the accuracy averaged within each participant and that among the participants. As a result, 80.6% of the WRITING-SOUND and 82.2% of the NON-WIRITING-SOUND were classified correctly on average.

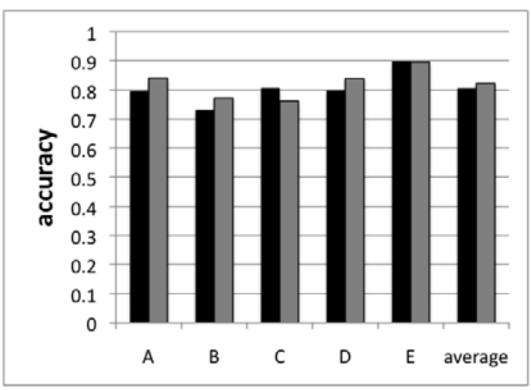


Fig. 6. The classification accuracy of vibration sound when the each five subjects marked a correct answer. The accuracy of the WRITING-SOUND is colored by black, and gray bar shows the accuracy of the NON-WRITING-SOUND.

[1] MVPen Technologies Ltd. (http://www.mvpen.com/)

4.2 Evaluation of Subjective Difficulty Estimation

We evaluated the proposed subjective difficulty estimation method described in Sec. 3.4 by conducting a user study. We prepared three math problems of different difficulties (linear equation, simultaneous equation, and quadratic equation). A subject tried to solve them in a random order. The subjective difficulty of each problem was computed by Eq.(1). We also asked the subject to report the subjective difficulty of each problem, according to a 5-point Likert scale from 1=*very easy* to 5=*very difficult*, just after he (or she) solved it.

Ten university students (6 male, 4 female) participated in the study. As a result, the order of the three problems according to the computed subjective difficulty and that to the reported subjective difficulty corresponded in six subjects' data. Figure 7 shows the computed and reported subjective difficulties of the problems of three subjects. Figure 7 (a) and (b) show the similar characteristic, while Fig. 7(c) shows the other type.

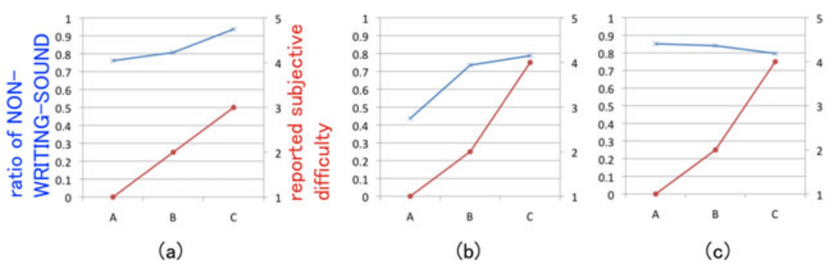


Fig. 7. The evaluation result of subjective difficulty estimation of three subjects. The line with cross dots shows the ratio of NON-WRITING-SOUND, and the line with circle dots shows reported subjective difficulty.

5 Discussion

It is regarded that one of the reasons why handwriting estimation accuracy was around 80% was the difference in writing pressure among the subjects. In the current implementation, the measured sound data was processed with a single decision tree which was constructed regardless of the subject's differences. Therefore the estimation accuracy would be degraded. Preparing different decision trees for different students would be the solution for this problem.

From the result of the evaluation conducted in Sec. 4.2, we confirmed that the correct subjective difficulty could be estimated for 60% of the subjects. Although there might be several reasons for the low accuracy, we think that this is mainly caused by the difference in behavior among individual subjects and by uncertainty of reported subjective difficulty as true value.

We employed the ratio of the duration of the NON-WRITING-SOUND to the total amount of time spent for a problem as the subjective difficulty, but this measure might not be universally applied. The behavior of a student who is working through a difficult problem is different among individuals. When deeply thinking about the problem, some students would stop moving their pens while others would tap their pens on the desk frequently. We designed the current system by taking into account the former case, so the later students are not supported. We will further investigate the behavior of students when they are working through difficult problems, and apply other distinctive features to estimate the subjective difficulty.

It takes a lot of time to develop an algorithm for individual identification because measurement and analysis of a great amount of data are needed in advance. Once the algorithm is developed, it is regarded that this technology could be efficient for application and the possibility of application would be widen. For example, a personalized study support will be available as we mentioned in Sec. 1. With the support system, it is also possible that controlling difficulty level of a problem and suggesting heavy review on the point where a student does not understand enough.

Furthermore, it is regarded that physiological measurement is significant to get the subjective emotion. It is generally said that emotion should be evaluated by the three measurements holistically; physiological, subjective, and behavioral measurement. Our research evaluates only subjective measurements reported by questionnaire. This seems to have a problem, the emotion which occurs for a moment on while the student is working through a problem might not be reflected in the questionnaire. This is because emotion is transient and is not memorized all of the time. If we apply the physiological measurement, the estimation accuracy would improve and we would be able to get hints for estimating other emotions because of obtaining the unconscious signal which could miss reported from questionnaire. However, a special equipment should be prepared to obtain physiological measurement. It is regraded that such measurement usually let subjects feel a kind of restriction because of sensor the user wears occurs a subject an extra stress which influences to his or her real studying. Thus we do not apply the physiological measurement in an ambient way.

To measure subjective emotion is supposed to obtain physiological measurement such as brain wave, it is difficult to apply it when students study in their classroom or home in general. A teacher also may not be able to observe all his students all the time. Therefor, we regard that the potential of the proposed approach is to estimate subjective difficulty by observing student's study all the time with only one microphone. This sensing environment is simple so that students must not feel any extra stress and could keep their natural study.

6 Conclusions

In this paper, we proposed a system which estimates the subjective difficulty of the student. First, our system obtained the vibration sound on the desk and analyzes the sound data. The student's behavior was classified into the WRITING or NON-WRITING categories. At last, the subjective difficulty was estimated by the ratio of NON-WRITING-SOUND.

As for future work, we will try to estimate not only subjective difficulty but also other subjective information such as uncertainly for a question. We will also apply other sensors to estimate the subjective difficulty more correctly.

References

1. Kappor, A., Burleson, W., Picard, R.W.: Automatic prediction of frustration. International Journal of Human-Computer Studies 65(8), 724–736 (2007)
2. Gunes, H., Piccardi, M.: Bi-Modal Emotion Recognition from Expressive Face and Body Gestures. Journal of Network and Computer Applications 30(4), 1334–1345 (2007)
3. Okubo, M., Fujimura, A.: Development of Estimation System for Concentrate Situation Using Acceleration Sensor. In: Human-Computer Interaction. New Trends, vol. 5610, pp. 131–140 (2009)
4. Lv, H., Lin, Z., Yin, W., Dong, J.: Emotion recognition based on pressure sensor keyboards. In: International Conference on Multimedia and Expo, pp. 1089–1092 (2008)
5. Luria, G., Rosenblum, S.: Comparing the handwriting behaviours of true and false writing with computerized handwriting measures. Applied Cognitive Psychology (2009), doi:10.1002/acp.1621
6. Harrison, C., Hudson, S.E.: Scratch Input: Creating Large, Inexpensive, Unpowered and Mobile Finger Input Surfaces. In: Symposium on User Interface Software and Technology, pp. 205–208 (2008)
7. Nakamura, K., Kakusho, K., Murakami, M., Minoh, M.: Estimating Learners' Subjective Impressions of the Difficulty of Course Materials in e-Learning Environments. In: APRU 9th Distance Learning and the Internet Conference 2008, pp. 199–206 (2008)

Ontology Driven Piecemeal Development of Smart Spaces

Eila Ovaska

VTT Technical Research Centre of Finland, Oulu, Finland
Eila.Ovaska@vtt.fi

Abstract. Software development is facing new challenges due to transformation from product based software engineering towards integration and collaboration based software engineering that embodies high degree of dynamism both at design time and run time. Short time-to-markets require cost reduction by maximizing software reuse; openness for new innovations presumes a flexible innovation platform and agile software development; and user satisfaction assumes high quality in a situation based manner. How to deal with these contradictory requirements in software engineering? The main contribution of this paper is a novel approach that is influenced by business innovation, human centered design, model driven development and ontology oriented design. The approach is called Ontology driven Piecemeal Software Engineering (OPSE). OPSE facilitates incremental software development based on software pieces that follow the design principles defined by means of ontologies. Its key elements are abstraction, aggregation and adaptivity. The approach is intended for and applied to the development of smart spaces.

Keywords: smart space, ontology, context awareness, interoperability, MDD.

1 Introduction

In our every day life we exploit a great number of software services. These services are provided through the Internet, devices and systems embedded into our surroundings. Our aim with smart spaces is simple [1]; to make information about the surrounding physical environment easily available, and thereby encourage (end-) users to make innovative applications by combining and using freely available information for new purposes.

A variety of terms is used to define the nature of a smart space. In [2], a smart space is defined as computational intelligence distributed and embedded into users' surroundings, including, for example, sensors and home appliances connected to each other via wired and wireless networks. This definition is equivalent to our definition of a physical space that is related to the physical surroundings of the space users, e.g. home, office or city. On the contrary, a global space focuses on making information available for global users. Thus, the spatial property of a space is different. Moreover, a logical space is an entity that combines a set of actors in a predefined way (e.g. a meeting) or spontaneously for a specific purpose (e.g. a chat in a social community). Although these definitions of spaces are different, one property is common for all of them; the dynamic nature of spaces. For example, i) a space can be established in an

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 148–156, 2010.

ad hoc manner; ii) services (and their execution platforms) are emerging and disappearing; and iii) services are used by different ways and in different contexts. This kind of dynamism is not possible to describe by using existing model driven engineering methodologies. Therefore, a new way of developing smart spaces is needed.

The contribution of this paper is an Ontology driven Piecemeal Software Engineering approach that exploits and enhances a set of existing approaches. The core is the assets repository [3], called smart space knowledge base that embodies knowledge in the form of ontologies, exploited in a set of software engineering phases; i) business and collaboration based innovation inspired by software product line engineering [5]; ii) scenario based design that brings user's viewpoint to the design [6]; iii) context-aware architecting based on model driven engineering and semantics modelling [7]; and iv) scenario based evaluation and testing based on [4]. Moreover, the (smart space) development follows SCRUM [8] that provides advantages for the incremental and integration oriented software development.

The structure of this paper is as follows. Section 2 presents background information. Section 3 introduces the OPSE approach and exemplifies its usage by smart space development. Concluding remarks close the paper.

2 Background

Some new model-driven approaches have emerged based on the fact that Unified Modeling Language (UML) is the most widely accepted modeling language. Typically these approaches introduce a meta-model enriched with context related artifacts, in order to support context-aware service engineering. For example, Achilleos et al. [9] present a context-aware pervasive service creation framework including a context ontology, a context modeling language and a tool environment that assists in context-aware service creation. The approach tackles the structural and static part of pervasive service creation but dynamic aspects of context-aware services are not supported. Moreover, Soylu et al. [10] introduce an approach that exploits the benefits of model driven and ontology driven development. The integrated process model is abstract and does not explicitly specify how the context ontology is to be specified, represented, and processed at design time and at run time, and how the context ontology is to be transformed for the use of different architectural elements, i.e. applications, services and data.

Our earlier work has focussed on enhancing the use of architectural knowledge in the development of service oriented architectures. First, we defined a service taxonomy, a reference service architecture and a set of basic services for developing wireless software services by exploiting existing styles, patterns and standards for providing artefacts of high quality [3]. Second, we enhanced our approach by defining a method for specifying quality requirements in an unambiguous way [11] so that systematic quality evaluation, e.g. integrability, extensibility and reliability evaluation could be carried out. Moreover, we have developed a set of tools that facilitates the use of the developed methods, as shown in [4]. Recently, we have studied run-time quality management [13] and context-awareness [14] that is a prerequisite for run-time quality management. One obstacle is the immaturity of context ontologies proposed for pervasive computing environment; none of them have been standardized nor widely accepted and systematically used. Typically, one of the existing context ontologies is used as a starting point and enhanced for the needs of the application domain in question. Moreover, context reasoning is poorly supported; some software technologies based on a specific language like SWRL or a micro architecture such as the Event-Condition-Action concept have been introduced but they are not widely accepted due to lack of maturity and tool support.

3 Piecemeal Development Based on Knowledge Base

The goal of OPSE is to manage evolution by abstraction, to assist integration oriented software development by aggregating patterns and ontologies of business, collaboration, and software development, and to create smart applications based on adaptive execution platform (Fig. 1).

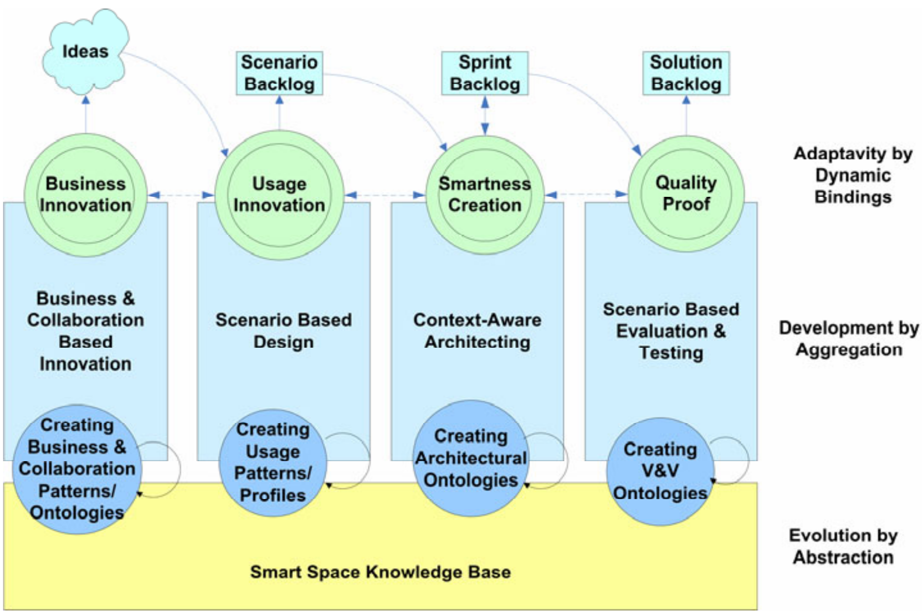


Fig. 1. An overview of the ODPD approach

The approach has two paths:

- To develop an ontology driven (architecture) knowledge base that helps in each development phase of smart spaces. The main contribution of this path is a set of ontologies and reusable patterns that are used for developing an execution platform of the smart space taking into account the needs of business, collaboration, users, development and operation.
- To provide an easy application development approach that encourages (end-)users to create innovative solutions on top of the smart space execution platform.

Development of smart spaces has the following phases; business and collaboration based innovation; scenario based design; context-aware architecting and scenario based evaluation and testing. The name of each phase illustrates the goal and the used method of the phase. Thus, the first phase focuses on business and collaboration. The driver of the second phase is users' satisfaction. Architecting solves how situation based behaviour is to be realized. And, finally scenario based evaluation guarantees that users' satisfaction is achieved and the business and collaboration rules are followed.

Each phase provides a piece of knowledge as ontology to the Smart Space Knowledge Base and a set of principles that define how knowledge is to be applied in software development:

- business, collaboration and usage patterns/ontologies/profiles are used for configuring and adapting the behaviour of the building blocks of smart spaces;
- architectural ontologies are used for defining, measuring, reasoning and adapting the functional and quality properties of smart spaces;
- innovation is encouraged by easy-to-use tools; executed on top of the smart space knowledge base and applied to smart space application creation and testing; and
- creation of smartness is made to be happened by making it easy to smash up building blocks at design-time and run-time, as a result of reasoning embodied in the space and its applications.

Typically architectural views are used for different purposes, e.g. for describing structure, behaviour and deployment. Piecemeal software development means that pieces are made by different stakeholders, not only by architects. Pieces may be different kinds of artefacts; ontologies, models, patterns, metrics, guidelines and source code. Thus, pieces do not replace views but facilitate the knowledge creation and usage in an incremental and long lasting development and evolution of smart spaces. Moreover, the piecemeal platform development focuses on creation and management of the smart space knowledge base, whereas the smart space application development exploits that knowledge in incremental application development by exploiting the main principles of SCRUM.

3.1 Piecemeal Platform Development

Business and collaboration patterns and ontologies. The purpose of the intended smart space is defined from business and organisation points of view by answering to the following questions; Is the space to be profitable? Is the main goal to provide enhanced experiences to its users or something else? How to estimate business impact, markets and acceptability of the smart space? What is the value creation model? What are the stakeholders' business models and roles in the value creation model? How does the smart space relate to the stakeholders' other businesses? What are the stakeholders' roles and responsibilities in smart space initiation and operation?

The above mentioned questions are answered by defining a set of patterns that define the viable options. When these patterns are validated in practice they are described as a business ontology and collaboration ontology used as part of the smart space operative management. The concepts, properties and rules of the defined ontologies share the managerial knowledge in smart spaces and keep the consistence of decisions made in a dynamic business and collaboration environment.

So far, we have developed a set of smart spaces where multi-vendor devices interoperate perfectly together. We have identified a set of collaboration patterns but creation of business patterns is still pending. Next, we focus on exploring multi-domain smart spaces, e.g. how to harmonize business/collaboration between a personal space, smart home and smart city. By a cross-domain experiment we try to identify common business/collaboration patterns used in all domains or part of them.

Usage patterns and profiles. User orientation is the obligatory feature of all smart spaces. Therefore, user centric design of smart spaces benefits from common reusable or adaptable usage patterns and profiles. The definition of these patterns/profiles is based on the results achieved from demonstrations that are analysed and thereafter, described and refined as reusable usage patterns and profiles. Patterns are used as building blocks in smart application development (i.e. smartness creation in Fig. 1) for adapting and configuring a generic application model for a particular usage scenario. Profiles are used for run-time configuration and adaptation with domain or usage specific properties and for evaluating and testing usage scenarios as shown in [4].

So far, we have developed patterns for role based adaptation, security based adaptation, and usage profiles for configuration and evaluation. We assume that the development of smart spaces will widely benefit from usage patterns and profiles at development time and run time due to the user centric design and context-aware execution that are crucial properties of smart spaces.

Architectural ontologies. Smart spaces embody service orientation and dynamism that can be triggered by any change in the user's intent, external or internal behaviour of the space. Therefore, the architectural ontologies consists of the ontologies that define the functional and quality properties of services and data, metrics and measuring techniques of the quality attributes that need to be managed and the context ontology at three levels; physical, digital and situational [14]. The physical level defines context related information available from physical entities. The digital level describes which kind of context information is available from digital entities, e.g. services and applications. Digital context is, in many cases, a fusion of a set of physical context information. Situation context refers to the situation of the application provided by the smart space. The situation context is the most abstract representation of the fusion of physical and digital context information and used together with the user context information for adapting the application for the needs of the end-user defined as a profile of preferences.

The layered abstractions of ontologies define the concepts and properties of services and data. First, the foundational ontology specifies a set of generic concepts that are common across domains, i.e. problem independent specifications. Second, the core ontology provides a top-level abstraction of concepts and properties following the CONON ontology by defining four clusters [15]; endurant (physical, non-physical, capability), perdurant (state-based, event-based), characteristics (physical, temporal, abstract) and abstract entities without spatial and temporal properties. Third, domain ontologies define domain specific concepts derived from the large set of scenarios defined for each domain and the earlier developed domain ontologies. In large smart spaces, e.g. a smart city, also application specific ontologies may be needed for making agents to interoperate with each other.

Quality attribute ontology may include different things; objectives, assets, mechanism, metrics, measuring techniques, etc. We have selected information security into our focus, explored the existing security ontologies and identified that all the other characteristics except metrics and measuring techniques are already covered quite well by the existing ontologies. Thus, we focus on security metrics and measuring techniques, keeping as a starting point the security ontology defined in [4].

So far, we have defined a micro architecture that includes the measuring, analysis, reasoning and adaptation activities that exploit the context and security ontologies for making run-time adaptation of the space's behaviour according to the quality goals, identified threats and available security solutions [16].

V&V ontologies. V&V ontologies with supporting environment can be developed based on the knowledge on the usage patterns and profiles, metrics and measuring techniques defined for run-time monitoring, and the existing evaluation and testing techniques. After the first experiences of the development of smart spaces, we decided to study further how to exploit the earlier developed quality evaluation and testing methods, e.g. in [4], the recently developed visualization tool [17] used for visualizing structure, behavior and quality attribute variation in smart spaces and the defined usage patterns and profiles in order to automate part of testing activities. This work has recently started.

3.2 Piecemeal Application Development

Business innovation exploits the knowledge described as business and collaboration patterns and ontologies in the platform development. However, if this kind of business-collaboration knowledge base is not available, as it was in our case when we started our first smart space developments, the following activities were carried out, which resulted in a set of ideas of the types of smart spaces:

- *Scoping a smart space* defines the purpose and boundaries of the space from owners, users, provided services, importance, quality and risks points of view.
- *Impact analysis* deals with what kinds of values, interests groups, and market penetration the smart space will provide.
- *Risk analysis* covers business and technical risks related to the objective of the space, the roles of partners involved in the smart space development, return on investments made on technology and skills development in a specific time period.
- *Assets analysis* includes the value analysis of potential existing software and hardware solutions and open source, which could be exploited in smart space development. The analysis covers functionality, quality, constraints and potential of existing assets in the smart space development.

Innovative usage scenarios. Eliciting usage scenarios is the most important activity in smart space application development. Brainstorming and team work were used as means of mining new innovative applications. Results were described by using a particular scenario template that starts with a short narrative description and continues with the intent, use cases, stakeholders, shared information, actors, artefacts, quality properties, triggers etc. The actors are potential agents (i.e. cooperating software entities) used further for breaking a scenario into sprints. A sprint consists of a set of software features that have to be developed in a certain period of time, e.g. in twenty days.

Prioritization and selection of the usage scenarios result in a backlog of prioritized scenarios, each of which describes a smart space application that is estimated to have high potential, business impact and market penetration in a reasonable time frame. Prioritization helps in selection which scenarios are beneficial to implement and in which order. Thus, this prioritized list of scenarios builds up the piecemeal design and implementation of applications.

Smart space applications. What makes smart spaces 'smart'? It is a space that facilitates semantic interoperability of applications in a situation based manner. Thus, the platform has to provide application interoperability at the following levels (in the bottom-up order); connectivity; network; data; meaning of data; context; change of context; and behaviour. We defined an architectural style that focuses on information interoperability (data, meaning of data, context and meaning of context) and tolerates heterogeneity in the lower interoperability levels (connectivity and network) [18]. The style embodies sixteen principles related to shared information, simplicity, agnostics, and extensibility among others. In summary, these principles define the building blocks required for facilitating piecemeal development of smart space applications.

Context-awareness is also facilitated by the context ontology and specific services provided by the smart space platform [14, 19]. Behavioural interoperability is achieved by the mechanisms that take care of context-based reasoning, learning and execution of adaptation actions. So far, simple reasoning mechanisms have been developed and tested in demonstrators but our intent is to create a set of sophisticated algorithms that are applicable to a set of problems common to all smart spaces. Thus, these concepts still need further analysis, design, implementation and testing in real smart environments.

Model driven application development is supported by a tool that assists in the use of the defined ontologies with the domain specific application modelling language, supported by the SmartModeller tool [20]. Currently, Java and Python implementations can be generated from designs. However, our goal is to enhance the tool to support also programming languages, commonly used in embedded devices, namely C and C++.

Demonstrations. All implemented scenarios are tested in their real environments by exploiting the existing assets, i.e. devices, sensors, networks, systems and software. The quality evaluation and testing is also incremental; First, each agent is implemented and tested separately (functional testing). Second, the agent is tested with the platform (interoperability testing). Third, the agent is tested with other agents that form the application (collaboration testing). The demo video illustrates how a smart space is created and enhanced with applications [21]. The visualisation tool [17] was developed for monitoring that the desired quality levels are met and validating that the implemented run-time adaptation rules work as intended. Finally, the smart space application (and the space itself) is ready for a field test. A set of field tests have been planned but it takes time to get feedback from these tests.

4 Conclusions

Software engineering is facing new challenges in the development of dynamic, context-aware and interoperable smart spaces that have to provide multi-domain assistance optimized to the user situation. In this paper, a novel ontology driven piecemeal development approach, OPSE, was introduced. The idea behind the approach is to develop a smart space incrementally, based on pieces of artefacts and software, by utilizing ontology oriented design for creating smart space knowledge base used as a knowledge management platform for agile application development. The use and applicability of the approach are justified by illustrative examples and references to a set of publications that explain the introduced concepts, models and solutions in detail. As a summary, we have demonstrated that the approach works in practice on the European level, where project partners form a heterogeneous team with multiple vendors and technologies. In the future, we intend to demonstrate that our approach works in heterogeneous multi-vendor and multi-domain settings.

References

1. Sofia/Artemis project, `http://www.sofia-project.eu/`
2. Okoshi, T., Wakayama, S., Sugita, Y., Aoki, S., Iwamoto, T., Nakazawa, J., Nagata, T., Furusaka, D., Iwai, M., Kusumoto, A., Harashima, N., Yura, J., Nishio, N., Tobe, Y., Ikeda, Y., Tokuda, H.: Smart space laboratory project: Toward the next generation computing environment. In: International Workshop on Networked Appliances (2001)
3. Niemelä, E., Kalaoja, J., Lago, P.: Towards an Architectural Knowledge Base for Wireless Service Engineering. IEEE Trans. on Software Engineering 31(5), 361–379 (2005)
4. Ovaska, E., Evesti, A., Henttonen, K., Palviainen, M., Aho, P.: Knowledge based quality-driven architecture design and evaluation. Information and Software Technology 52(6), 577–601 (2010)
5. Pohl, K., Böckle, G., van der Linden, F.: Software Product Line Engineering. In: Foundations, Principles, and Techniques, p. 466. Springer, Heidelberg (2005)
6. Ikonen, V.: Scenarios in Ubiquitous Computing System Design: User-driven vs. Technology-driven Usages. In: The 11th International Conference on Human-Computer Interaction (HCII 2005). Lawrence Erlbaum Associates, Mahwah (2005)
7. Bettini, C., Bridiczka, O., Henricksen, K., Indulska, J., Niclas, D., Ranganathan, A., Riboni, D.: A survey of context modelling and reasoning techniques. In: Pervasive Mobile Computing, vol. 6, pp. 161–180 (2010)
8. Scrum: What is Scrum? Retrieved March 8 (2010), `http://www.scrumalliance.org/learn_about_scrum`
9. Achilleos, A., Yang, K., Georgalas, N.: Context modelling and a context-aware framework for pervasive service creation: A model-driven approach. In: Pervasive and Mobile Computing, vol. 6, pp. 281–296 (2010)
10. Soylu, A., De Causmaecker, P., Desmet, P.: Context and Adaptivity in Pervasive Computing Environments; Links with Software Engineering and Ontological Engineering. Journal of Software 4(9), 992–1013 (2009)
11. Niemelä, E., Immonen, A.: Capturing quality requirements of product family architecture. Information and Software Technology 49(11-12), 1107–1120 (2007)

12. Chen, H., Finin, T., Joshi, A.: The SOUPA Ontology for Pervasive Computing. Whitestein Series in Software Agent Technologies. Springer, Heidelberg (2005)
13. Niemelä, E., Evesti, A., Savolainen, P.: Modeling Quality Attribute Variability. In: 3rd Int. Conf. on Evaluation of Novel Approaches of Software Engineering, ENASE 2008, pp. 169–176 (2008)
14. Pantsar-Syväniemi, S., Simula, K., Ovaska, E.: Context-awareness in Smart Spaces. In: The 1st International Workshop on Semantic Interoperability for Smart Spaces, ISCC 2010, Riccione, Italy, June 23, pp. 1023–1028 (2010)
15. Gu, T., Wang, X.H., Pung, H.K., Zhang, D.Q.: An Ontology-based Context Model in Intelligent Environments. Connected Home Lab, Institute for Infocomm Research, Singapore
16. Evesti, A., Pantsar-Syväniemi, S.: Towards micro architecture for security adaptation. In: 1st International workshop on Measurability of Security in Software Architectures, ECSA 2010, Copenhagen, Denmark, August 23, pp. 181–188 (2010)
17. Kuusijärvi, J., Evesti, A., Ovaska, E.: Visualizing Structure and Quality Properties of Smart Spaces. In: 1st International Workshop on Semantic Interoperability for Smart Spaces, ISCC 2010 (2010)
18. Ovaska, E., Salmon Cinotti, T., Toninelli, A.: Design Principles and Practices of Interoperable Smart Spaces. In: Xiaodong, Yang (eds.) Advanced Design Approaches for Emerging Software Systems, IGI Global Research Handbook, p. 30 (to appear)
19. Toninelli, A., Pantsar-Syväniemi, S., Bellavista, P., Ovaska, E.: Supporting Context Awareness in Smart Environments: a Scalable Approach to Information Interoperability. In: Proc. of the Int. Workshop on Middleware for Pervasive Mobile and Embedded Computing, SESSION: Short papers, Article No: 5. ACM, IFIP, USENIX (2009)
20. Katasonov, A., Palviainen, M.: Towards ontology-driven development of applications for smart environments. In: Workshop on the Web of Things, PerCom 2010, pp. 696–701. IEEE, Los Alamitos (2010)
21. Semantic Interoperability in Smart Spaces. The Green House Demonstration video, http://www.youtube.com/watch?v=EU9alk9t7d

Exploiting Acoustic Source Localization for Context Classification in Smart Environments

Christian Kleine-Cosack, Marius H. Hennecke,
Szilárd Vajda, and Gernot A. Fink

Robotic Research Institute, TU Dortmund University,
Dortmund, Germany

Abstract. Smart environments rely on context classification in order to be able to support users in their daily lives. Therefore, measurements provided by sensors distributed throughout the environment are analyzed. A main drawback of the solutions proposed so far is that the type of sensors and their placement often needs to be specifically adjusted to the problem addressed. Instead, we propose to perform context classification based on the analysis of acoustic events, which can be observed using arrays of microphones. Consequently, the sensor setup can be kept rather general and a wide range of contexts can be discriminated. In an experimental evaluation within a smart conference room we demonstrate the advantages of our new approach.

Keywords: Context classification, audio localization, smart rooms.

1 Introduction

With the advent of pervasive computing technologies, so-called smart spaces have been introduced to support humans during their activities of daily living (ADL). Varieties of sensors are embedded into the people's homes and working environments constantly monitoring activities and surrounding conditions. To allow for practical application, the sensor setups must be chosen with care, respecting the specific requirements of the scenario. In the domestic domain, e.g., privacy issues are one key concern, while the flexibility of the sensor setup is another for smart spaces which are used in a wide range of application. Typical use cases for the latter are smart conference rooms and working environments, which support the human users during meetings and presentations or assist in collaborative working tasks (e.g. [16]).

A fundamental prerequisite for a number of applications related to smart spaces is the precise knowledge of the environment's context. In a smart conference room, e.g., intelligent controlling of light sources depends on knowledge about the current situation. While the luminosity must be kept below a certain level in the projection area during presentation, the opposite is desired for discussions and collaborative tasks. On a more abstract level, an important prerequisite for pro-active behavior of the smart environment is the knowledge about the current situation.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 157–166, 2010.

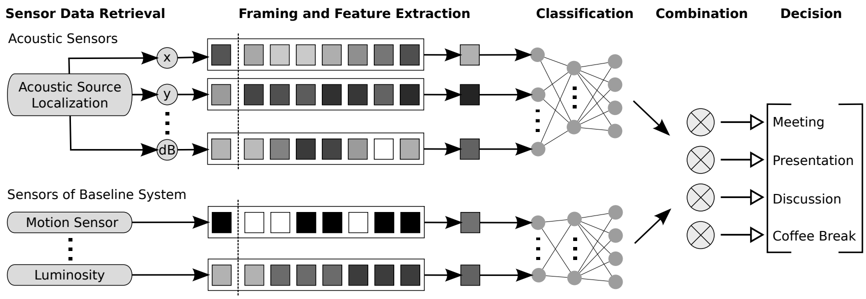


Fig. 1. System overview

In this paper we propose a real-time context classification system based on acoustic source localization (ASL). As smart conference rooms are usually equipped with multiple microphones, acoustic events within the environment can be constantly monitored. The configuration of the microphone setup is chosen w.r.t. the characteristics of this sensor type only, not considering the specific experimental setting. This generic setup allows the classification of a multitude of situations, without the need to adopt the sensors' configuration to the latter. Hence, it provides high flexibility in practical applications. For reasons of comparison and to further improve the performance, a baseline system utilizing environmental sensors (i.e. motion, air condition, luminosity and door status sensors) is used and integrated into the classification process.

For classification, the signal energy as well as the spatial position of acoustic events are calculated for each time frame and statistical features are derived, i.e. mean and variance. Using a sliding window approach the latter are accumulated over time, and the resulting vector is fed into a Multi-Layer Perceptron (MLP) based classification system. An overview of the system structure is given in Fig. 1 illustrating the ASL based classification, the baseline system, and the integration of the two. Direct measurements of the environmental sensors are used in the feature extraction step whereas an acoustic source localization is used beforehand for the acoustic sensors. The sensor types are treated separately in means of classification and a final combination step results in a context decision.

The effectiveness of the given approach is demonstrated in an experimental evaluation in a smart conference room. In the given scenario the results show improved classification rates compared to the baseline system. Moreover, we show that the two can be combined to improve the overall performance. The proposed system is the first that allows context classification in smart environments based on the acoustic analysis of human interaction patterns. The sensor configuration is completely independent of the situation to classify in a specific scenario, hence, it provides high flexibility in practical applications. Moreover, only statistical data is used for the classification of context. In consequence, the presented approach can also be applied to scenarios in which privacy requirements must be met. The presented system is capable of real-time context classification, a key requirement in practical applications.

2 Related Work

A large volume of literature on situation analysis in smart environments exists, which can be split into activity recognition (AR) and context classification. While the former focuses on the analysis of ADL, the intention of the latter and the proposed approach is to recognize the overall context of the environment, i.e. the current, long-term situation of a smart room.

2.1 Situation Analysis in Smart Environments

The literature covers AR either based on computer vision [15] or by applying pervasive computing approaches [1]. Whereas AR systems are usually based on statistical modeling frameworks, such approaches can rarely be found in case of context classification.

The evaluation of ontologies represents the methodology of choice instead. An ontology represents a well-defined, expert-made set of rules, which describes situations in certain environments by means, e.g., of (fixed) combinations of sensor events. For the conceptual and technical implementation of ontologies different semantics and toolkits are used [6,14]. Also, semi-automatically constructed ontologies are described that allow to split the context model into a handcrafted and an automatically derived part [12]. Whereas the first part covers more general, domain-related aspects, the latter focuses on specialties of the particular application. Still, prior expert knowledge remains an integral part of these rule based approaches. In consequence, ontologies provide a methodology for context classification in scenarios which can be described by a set of simple rules. Such prerequisite can be met in case of sensor setups tailored towards a specific situation, in which a predefined combination or sequence of sensor events determines the current situation. In case of continuous, rather noisy and widespread sensor events; however, the application of ontologies is limited and more likely to fail.

Situational context classification using statistical models and common sensors is hardly covered by related work. Using binary motion sensors in [18] topological sensor networks are evaluated for this task. Focusing on energy saving in [5] multi-modal environmental data (monitoring air quality, water consumption, motion etc.) are analyzed using a hidden Markov model (HMM). In [8] measurements from custom-made binary state change sensors are analyzed using an HMM and Conditional Random Fields. A similar sensor network is considered in [17] for tracking and activity recognition using a particle filter. However, all these approaches rely on a multitude of sensors which must be installed and configured precisely to cover possible human activities. Hence, expert knowledge and a deep understanding of the users' behavior is mandatory. In contrast, acoustic sensors can be installed without such specific knowledge, enhancing flexibility for practical applications.

2.2 Acoustic Source Localization

The problem of localizing the spatial position of an acoustic event is known as acoustic source localization (ASL). Several techniques for solving this task using multiple acoustic sensors, i.e. microphones, are known. The main challenges for an ASL system in a real world environment are noise and reverberation. Possible noise sources are fans of electrical equipment, foot fall sounds or any other putative unwanted sound events. Robustness against these factors are of utmost importance, in order to obtain reliable results from an ASL system.

Time delay estimation based algorithms [3] first determine the delay of a signal between two or more sensors and combine those time-difference of arrivals (TDOA) in a second stage to form a source location. Not taking an explicit room reverberation model into account, the generalized cross-correlation (GCC) [10] provides a simple and efficient approach. Unlike the GCC method adaptive eigenvalue decomposition (AED) [2] models reverberation explicitly, by blindly estimating the acoustic channels impulse responses. Its higher computational costs are justified by a more robust TDOA estimation in reverberant environments. Each TDOA restricts the acoustic source to lie on one sheet of a two-sheeted hyperboloid and hence the second stage needs to find the intersection of all TDOA parameterized hyperboloids (e.g. [7]).

Steered response power (SRP) based ASL schemes use beamforming techniques to steer a passive sensor array to different locations. The position with the highest beamformer output power is then assumed to be the source location. Hence, SRP algorithms avoid the two stage approach of the TDOA methods. Calculating the SRP for all possible locations requires a high computational effort. Nevertheless, DiBiase et al. [4] showed that the SRP of a simple delay-and-sum beamformer is equivalent to a spatial combination of all pair-wise GCCs, which leads to a computationally tractable ASL algorithm.

3 Context Classification Using ASL

Today, there is a trend to equip smart conference rooms with multiple microphones for teleconferencing applications or meeting recordings, which facilitates the extraction of positions of interacting human speakers. In the near future, distributed ad-hoc microphone arrays, e.g. the combination of mobile phones and other embedded devices available, might fulfill this task. In the following, we present our system for context classification in a smart environment which is based on ASL information.

3.1 Acoustic Source Localization

Steered response power based ASL methods have shown good results under moderate noise and reverberation levels. Due to this robustness and their additional simplicity, we employ such an SRP scheme—namely the SRP-PHAT [4]—and use the positions of the localized acoustic sources for context classification. Here, we will shortly revisit SRP-PHAT.

The TDOA of a source signal between two or more receivers is the basic building block of SRP-PHAT. A common approach for estimating a TDOA $\hat{\tau}_{ij}$ for the two-channel case is the GCC [10]

$$R_{ij}(\tau) = \mathcal{F}^{-1}\{\Psi_{ij}(\omega) X_i(\omega) X_j^*(\omega)\}(\tau) \ . \tag{1}$$

It is the inverse Fourier transformation $\mathcal{F}^{-1}\{\cdot\}(t)$ of the cross-spectrum of the acoustic channel i and j, where $x_i(t) = \mathcal{F}^{-1}\{X_i(\omega)\}(t)$ is the i-th channel's signal. In order to shape the GCC for better TDOA estimation performance the phase transformation (PHAT) [10] $\Psi_{ij}(\omega) = |X_i(\omega)X_j^*(\omega)|^{-1}$ is commonly used. It is motivated by the fact that a pure time delay results in a phase shift and leaves the signals amplitude unchanged. Hence, it is a simple whitening of the cross-spectrum. Despite its solely heuristic nature, PHAT as a pre-filter has shown robustness under moderate reverberation and noise conditions. For simplicity, we omit some technical details, e.g. a block-processing scheme.

The robustness of SRP-PHAT stems from the spatial combination

$$P(\boldsymbol{q}) = \sum_{(i,j)\in\mathcal{P}} R_{ij}(\tau_{ij}(\boldsymbol{q})) \ , \qquad \tau_{ij}(\boldsymbol{q}) = \frac{|\boldsymbol{q}-\boldsymbol{p}_i| - |\boldsymbol{q}-\boldsymbol{p}_j|}{c} \ , \tag{2}$$

of possibly redundant GCCs for pairs $(i,j) \in \mathcal{P} \subseteq \{1,\ldots,M\}^2$ using a total of M microphones. Given a spatial position $\boldsymbol{q}$ and the microphone positions $\boldsymbol{p}_i, \boldsymbol{p}_j$ the TDOA leading to $\boldsymbol{q}$ is calculated via $\tau_{ij}(\boldsymbol{q})$, where $c = 343\,\mathrm{m\,s^{-1}}$ is the speed of sound and assumed constant. Equation (2) defines a pseudo spatial likelihood function (SLF) for a source position $\boldsymbol{q}$. Hence, an estimate for the location of the dominant acoustic source is given as $\hat{\boldsymbol{q}} = \arg\max_{\boldsymbol{q}} P(\boldsymbol{q})$.

3.2 Feature Set

The aforementioned ASL system operates on frames of 300 ms. With a sampling rate of $f_s = 48\,\mathrm{kHz}$ and a frame shift of 150 ms the system estimates up to seven source positions per second. Each estimate includes the three-dimensional location $\hat{\boldsymbol{q}} = (x,y,z)^{\mathrm{T}}$, the SLF value $P(\hat{\boldsymbol{q}})$ and the speech energy of the current signal frame in dB averaged over all acoustic channels. For those five raw ASL measurements the mean and variance over two second windows (14 raw ASL samples) are calculated, leading to ten ASL features in total. The window size is chosen, such that a fair amount of context-relevant speech portions are considered.

Given those ASL features a sliding window procedure is applied, extracting frames of fixed lengths l with overlap $o = 10\,\mathrm{s}$. Elements of the frame are accumulated, resulting in a ten-dimensional vector which, after mean subtraction, is normalized element-wise and fed into the classification system.

3.3 Baseline System

In order to compare our approach for context classification based on ASL, we use a baseline system which utilizes only environmental, non-intrusive sensors [9]. Monitoring the temperature, CO_2 content and humidity of the air, luminosity, the status of doors and windows, as well as motion in certain regions of the environment, this system has already proven its capabilities for the given task. Moreover, this baseline system is integrated in the ASL based context classification approach to further improve the results.

3.4 Neural Network Classifier

In order to classify the smart environment's context based on audio source data (ASL approach) and sensory information (baseline system), a neural network has been built. Instead of using rule based ontologies which involves human expert knowledge, our system is based on statistical learning and can adapt to arbitrary data. For our experiments a fully connected multi-layer perceptron (MLP) with sigmoid transfer function as activation has been used [11]. For training purposes an error back-propagation mechanism has been employed to minimize a square-error metric.

The network topologies are directly derived from the data. The number of processing units in the input layer corresponds to the dimensionality of the data, while the output nodes are assigned to the different context scenarios to be classified. The neurons in the hidden layer can not be estimated automatically, hence some trial runs have been conducted in that sense. Considering the audio data, a 10-10-4, while for the environmental sensor data a 15-10-4 topology is applied. In order to integrate the two classifiers, first the activations of the output neurons of each network are normalized, i.e. estimates of class posterior probabilities are obtained. Assuming class independence between the different sensor information, combined estimates can be calculated by simple multiplication (right part of Fig. 1).

4 Evaluation

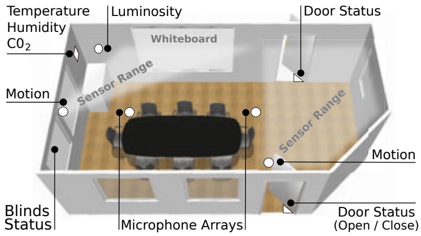
In order to demonstrate the effectiveness of the presented approach we conducted a number of practical experiments in a smart conference room. The objective of the experiments was to evaluate the given approach in a real world setting during regular usage of this environment.

4.1 Setting

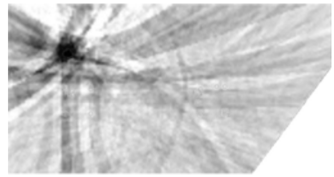
The work described in this paper is part of a greater research project, the FINCA [13]. This intelligent environment, amongst others, consists of a conference room equipped with a multitude of ambient sensors, microphones and cameras. Figure 2a gives an overview of the sensor setup in general. Sixteen microphones grouped in two distributed ceiling-mounted arrays are employed for ASL. A visualization of the SLF (2) for an arbitrary time instance as used for the ASL approach is shown in Fig. 2b. The location of the audio source is approximately the center of the high energy area (dark spot in the two-dimensional map).

4.2 Data Set and Methodology

Sensor data was recorded over a time period of several months with multiple recording sessions for each situation. The final data set consists of four situations that are typically related to meeting rooms.



(a) Configuration of acoustic and ambient sensors.



(b) SLF $P(\boldsymbol{q})$ evaluated over the two-dimensional FINCA grid.

Fig. 2. The smart conference room of the FINCA

Meeting: Team meeting, using possibly the white-board for collaborative tasks (122 minutes of data)

Presentation: A talk given by one person while slides are projected on the white-board. (107 minutes of data)

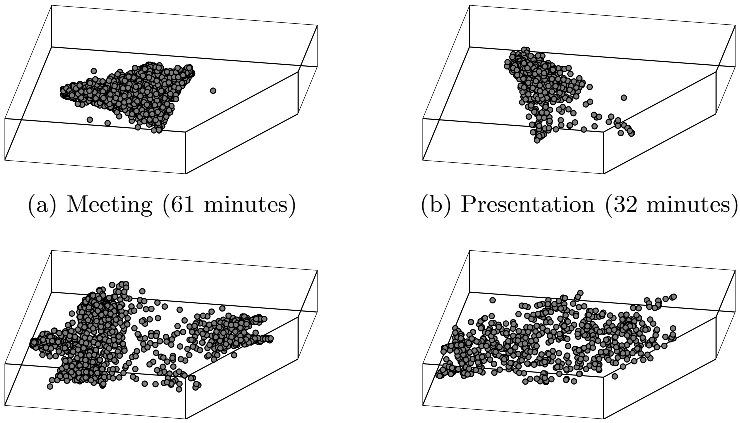
Discussion: Questions from the audience after a given presentation. (67 minutes of data)

Coffee break: Coffee break between subsequent presentations and meetings. (45 minutes of data)

Regarding the acoustic sensor information, separating these classes is a challenging task as audio events occur rather widespread throughout the different situations. Figure 3 visualizes the localization results for exemplarily selected, complete recording sessions. Note that for the classification only short time periods of such sensor data, i.e. several seconds, are used. Moreover, not the raw coordinates as plotted in Fig. 3, but the more general statistics are utilized. Hence, not the positions of speakers are learned by the classification system, but the patterns of interaction between multiple persons.

A five-fold cross-validation was performed for training, parameter optimization and final testing. The dataset was subdivided into five disjoint sets. Training was performed using three fifths of the sets, validation of the MLP on the fourth, and final testing on the remaining set. Final results were obtained by averaging over the results achieved for every such configuration. To account for practical applicability, no effort was taken to manually adjust the data set. For training, classes were automatically balanced by data duplication, whereas the validation and test procedure was run on the unmodified and unbalanced set.

The experiments were designed to analyze the classification performance for ASL based context classification, the baseline system and the combined approach for different window sizes. Moreover, a reduced feature set was tested to investigate the ability of the system to learn interaction patterns apart from any spatial position. As stated above, only derived statistics are used by the classifier, not the raw speaker locations. However, the mean value of x,y and z-coordinates over all audio events within a specific frame – i.e. the center of audio activity in



(a) Meeting (61 minutes) (b) Presentation (32 minutes)

(c) Discussion (46 minutes) (d) Coffee Break (22 minutes)

Fig. 3. ASL results for different exemplarily selected audio recordings

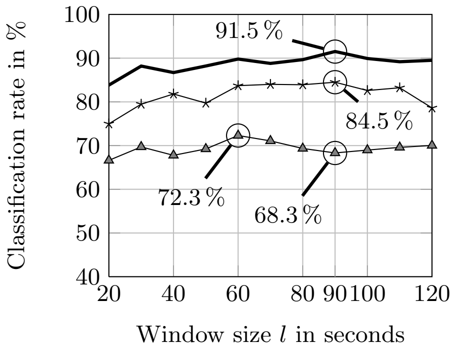
the environment at the given time – is regarded. In contrast, the reduced feature set discards this location specific information.

4.3 Results

Figure 4 (left part) shows the classification results for the different classification systems and window sizes. The ASL based approach outperforms the baseline system and reaches a classification rate of 84.5 % at window size $l = 90$ seconds, while the others best performance is 72.3 % at $l = 60$ seconds. Combining the two further improves the performance and the overall best of 91.5 % is achieved for $l = 90$ seconds. The corresponding confusion matrix and class accuracies are presented in Fig. 4 (right part). While the classes *meeting* (M), *presentation* (P) and *coffee break* (C) are classified correctly with rates lager 90 % the class accuracy of *discussion* (D) only achieves 80.8 % caused by the confusion with the class *presentation*. This can be explained by the similarity between the two situations, i.e. the former presenter might still be the main speaker when answering questions of the audience in detail.

The effect of the window size can be explained with the increasing amount of context information encoded into a frame which positively affects the classification. However, the smoothing characteristic of the framing decreases the performance for larger l. The perfect parameter value depends on the feature set and differs between the ASL approach and the baseline system.

Results on the reduced feature set (discarding mean values of x and y-coordinates) in comparison to the former are illustrated in Fig. 5 (left part). While the ASL based classification decreases slightly (83.1 %), the combined performance remains on the same level. In contrast, additionally discarding the mean of z, dramatically reduces the classification rates (57.5 % ASL, 73.4 % combined)(Fig. 5, right part). These findings show the capability of our approach to



	M	P	D	C	Class Acc.
M	441	7	2	0	98.0 %
P	0	1157	42	0	96.5 %
D	0	138	582	0	80.8 %
C	0	64	62	1243	90.8 %
				Overall Acc.	91.5 %

Fig. 4. Classification results: ASL, baseline system, combination. The confusion matrix for the optimal combined result is shown on the right.

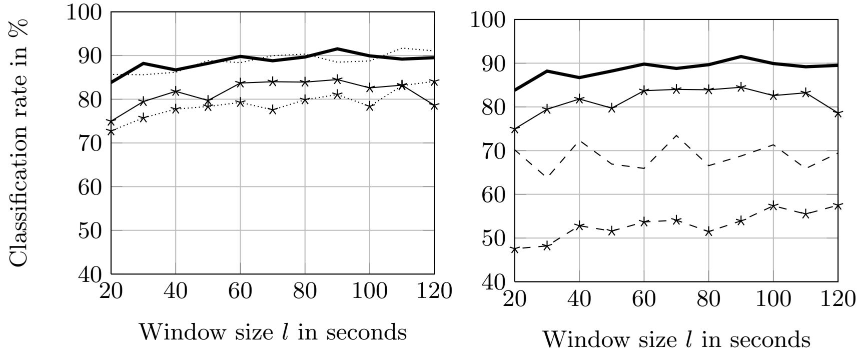


Fig. 5. Classification results for reduced ASL feature sets. Left part: Discarding mean of x, y and combined; Right part: Additionally discarding mean of z and combined. In comparison to the results presented in Fig. 4.

learn interaction patterns independent of the location of speakers. However, the spatial height of a localized audio source, e.g. whether different interaction partners are sitting or standing, has a great impact on the classification performance of our system and must be regarded for the given task.

5 Conclusion

In this paper we presented a system for context classification in smart environments. The proposed system is the first which utilizes acoustic information about human interaction patterns for this task. Without any prior knowledge a neural network classifier learns acoustic source localization patterns which are typical for specific contexts. Results of experiments in a smart conference room equipped with multiple microphones showed improved performance compared to a baseline system relying on environmental sensors. The integration of both information sources even increased the classification performance to 91.5 % in a demanding, real world scenario.

References

1. Atallah, L., Yang, G.Z.: The use of pervasive sensing for behaviour profiling - a survey. Pervasive Mob. Comput. 5(5), 447–464 (2009)
2. Benesty, J.: Adaptive eigenvalue decomposition algorithm for passive acoustic source localization. JASA 107(1), 384–391 (2000)
3. Chen, J., Benesty, J., Huang, Y.: Time delay estimation in room acoustic environments: An overview. EURASIP Journal on Applied Signal Processing (2006)
4. DiBiase, J.H., Silverman, H.F., Brandstein, M.S.: Robust localization in reverberant rooms. In: Microphone Arrays, ch. 8, pp. 157–180. Springer, Heidelberg (2001)
5. Dong, B., Andrews, B.: Sensor-based occupancy behavioral pattern recognition for energy and comfort management in intelligent buildings. In: Proc. Int. IBPSA Conf. (2009)
6. Henricksen, K., Indulska, J., Rakotonirainy, A.: Modeling context information in pervasive computing systems. In: Mattern, F., Naghshineh, M. (eds.) PERVASIVE 2002. LNCS, vol. 2414, pp. 167–180. Springer, Heidelberg (2002)
7. Huang, Y.A., Benesty, J. (eds.): Audio Signal Processing for Next-Generation Multimedia Communication Systems. Springer, US (2004)
8. van Kasteren, T., et al.: Accurate activity recognition in a home setting. In: Proc. UbiComp, pp. 1–9 (2008)
9. Kleine-Cosack, C., Plötz, T., Vajda, S., Fink, G.A.: Context classification in smart spaces using environmental sensors. In: AMIF Publication (to appear, 2010)
10. Knapp, C.H., Carter, G.C.: The generalized correlation method for estimation of time delay. IEEE Trans. Acoust., Speech, Signal Process 24(4), 320–327 (1976)
11. Kuncheva, L.I.: Combining Pattern Classifiers. Wiley-Interscience, Hoboken (2004)
12. Niu, W., Kay, J.: Location conflict resolution with an ontology. In: Proc. 6th Int. Conf. on Pervasive Computing, pp. 162–179 (2008)
13. Plötz, T.: The FINCA: A Flexible, Intelligent eNvironment with Computational Augmentation (2007), http://www.finca.irf.de
14. Ranganathan, A., Campbell, R.H.: A middleware for context-aware agents in ubiquitous computing environments. In: Proc. Int. Conf. on Middleware (2003)
15. Turaga, P., et al.: Machine recognition of human activities: A survey. IEEE Trans. on Circuits and Systems for Video Technology 18(11), 1473–1488 (2008)
16. Waibel, A., Schultz, T., Bett, M., Denecke, M., Malkin, R., Rogina, I., Stiefelhagen, R., Yang, J.: Smart: The smart meeting room task at ISL. In: Int. Conf. Acoustics, Speech and Signal Processing, pp. 752–755 (2003)
17. Wilson, D.H., Atkeson, C.: Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors. In: Pervasive Computing, pp. 62–79 (2005)
18. Wren, C.R., Tapia, E.M.: Toward scalable activity recognition for sensor networks. In: Hazas, M., Krumm, J., Strang, T. (eds.) LoCA 2006. LNCS, vol. 3987, pp. 168–185. Springer, Heidelberg (2006)

Real-Time Gaze Tracking for Public Displays*

Andreas Sippl[1], Clemens Holzmann[2], Doris Zachhuber[1], and Alois Ferscha[1]

[1] Johannes Kepler University Linz, Institute for Pervasive Computing
{andreas.sippl,doris.zachhuber,alois.ferscha}@jku.at
[2] Upper Austria University of Applied Sciences, Mobile Computing
clemens.holzmann@fh-hagenberg.at

Abstract. In this paper, we explore the real-time tracking of human gazes in front of large public displays. The aim of our work is to estimate at which area of a display one ore more people are looking at a time, independently from the distance and angle to the display as well as the height of the tracked people. Gaze tracking is relevant for a variety of purposes, including the automatic recognition of the user's focus of attention, or the control of interactive applications with gaze gestures. The scope of the present paper is on the former, and we show how gaze tracking can be used for implicit interaction in the pervasive advertising domain. We have developed a prototype for this purpose, which (i) uses an overhead mounted camera to distinguish four gaze areas on a large display, (ii) works for a wide range of positions in front of the display, and (iii) provides an estimation of the currently gazed quarters in real time. A detailed description of the prototype as well as the results of a user study with 12 participants, which show the recognition accuracy for different positions in front of the display, are presented.

1 Introduction

Eye gaze tracking, i.e. the measurement of a human's point of gaze, is a well-established method for evaluation purposes in the HCI domain [10], but it can also be used for implicit human computer interaction, which is defined as *"an action, performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as input"* in [9]. Similar to the interaction between humans, implicit HCI is very much based on contextual information, such as body postures, gestures, voice, eye movement and gaze. The knowledge of where people are looking at allows a system to react accordingly or to act pro-actively without explicit user input, e.g. by automatically adapting the information provided on a display to show more details or information of possible interest for the user.

Most non-body worn eye gaze tracking systems available are based on a stereo pair of cameras and the use of computer vision algorithms. Unfortunately, such systems are usually designed for (i) a single person only, (ii) computer screens

* This work is supported under the FFG Research Studios Austria program under grant agreement No. 818652 DISPLAYS (Pervasive Display Systems).

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 167–176, 2010.

up to 24 inch if the tracking system is mounted directly at the display, and (iii) the common reading distance to such a display. Applications with large-scale displays would require an individual hardware setup, like for example to install the tracking system in front of a display to cover its total surface [6], resulting in a higher setup and calibration effort as well as limitations in the freedom of body movement in front of the display due to the required minimum distance between the system and the display.

We have developed a new approach for gaze tracking which overcomes these limitations to enable multi-person usage for large displays in the public. For a first version, the system aims at distinguishing four areas from distances between one to four meters with a single camera. The underlying principle is to locate facial features like eyes, nose tip and the total face area, and to estimate the viewing direction from (i) the relative positions between these features as well as (ii) their positions with respect to the position of the detecting camera. Therefore, we do not track the eye gaze using e.g. the pupil-corneal reflection technique as described in [3], but rather estimate the head pose as surveyed in [5], from which we infer the user's point of gaze by assuming the eyes to be directed straight ahead. For many public display applications, we consider it sufficient to derive the focus of visual attention from the head pose, as the large display dimension forces the user to change the head orientation accordingly.

In this paper, we present (i) a content-adaptive pervasive advertising scenario, which makes use of gaze-based implicit interaction (Section 2), (ii) a literature overview about related work (Section 3), (iii) implementation details of our first prototype with the proposed gaze tracking approach (Section 4), and (iv) the design and results of a user study to evaluate its recognition accuracy for large-scale displays (Section 5). The paper is summarized in Section 6.

2 Application Scenario

A multi-user eye gaze tracking system for large-scale displays offers a wide range of applications, especially with respect to adapting the content to the estimated focus of visual attention. Our system is primarily designed for applications in the *pervasive advertising* domain, which explores the application of pervasive computing technologies for advertising purposes [1]. Pervasive computing environments allow the real-time delivery of relevant advertisements to selected consumers in suitable ways, thus reaching the right people with the right contents [8]. The visual focus of potential customers in front of public displays as well as an implicit kind of interaction are crucial for this purpose. Schmidt declares that implicit interaction is based on perception and interpretation [9]. In the pervasive advertising context, the perceived visual focus of a viewer can be mainly interpreted as interest in the presented product (next to pure visual stimuli by colorful or emotional imagery) – at least if the visual point of gaze persists over a certain amount of time. We have developed a system that aims at perceiving the visually focused region on a large-scale display, thus allowing the realization of advertising concepts for multiple users.

Interaction with a large-scale public display strongly depends on (i) the concurrent number of viewers and (ii) their distance to the content. According to [11], the area in front of a display can be divided into different interaction zones: *perception*, *view*, *read* and *touch*. They facilitate the transitions from implicit to explicit interaction as well as the transition from public to personal interaction depending on the distance to a display. Our proposed pervasive advertising concept addresses all of them, except for the explicit personal interaction by touch. In the advertising domain, we propose to use these zones based on our presented eye gaze tracking approach in the following way:

- *Perception Zone.* This state refers to the idle-mode of an advertising display, when no active viewers are detected. Generally, low-level information like (animated) graphics or text in large font size are displayed using the whole display area in order to gain the attention of passerbys.
- *View Zone.* When somebody is detected looking at the display at a medium distance of about two to four meters, the display area divides into sub-displays – e.g. four quadrants separating the area into up/down and left/right – showing similar information as the original content as well as new product categories. By analyzing which quadrant is observed by a user, the topic in which he is most interested in is detected. With this knowledge, more advertisements covering the same or similar product categories can be provided in the out-of-focus quadrants, to which he can switch if the content is of even more interest. This approach enables a subtle convergence to the most desired category. For multiple users however, more sophisticated techniques are required, but not described in more detail in this paper. Recommender systems, which run in the background and analyze previous user behavior to optimize the selection of offered categories [2], could improve the advertising effectiveness further.
- *Read Zone.* A similar adaptation approach is chosen at a close distance of less than two meters to the display. Instead of presenting similar advertisements, additional information to a product and details like text or graphical overlays are presented to the user, exclusively in the quadrant focused by the user. This is valid for single as well as for multiple users, as an extension of the current area of interest might be distracting, even if there is enough space available. The goal of this zone is to provide as much information as wanted and necessary to increase the chance of a purchase in the near future.

3 Related Work

There has been much work about gaze and head pose estimation already. Traditional eye gaze trackers have a high accuracy of about 1-3° and are mainly used as precise input devices, e.g. to control a mouse cursor. However, according to [3], a major problem of many traditional eye gaze trackers is their lack of important usability requirements, like for example being intrusive, requiring a long and complex setup and calibration procedure, or tolerating a very small head motion only. Newer approaches, so-called remote eye gaze trackers, are often based on the pupil-corneal reflection technique and tackle some drawbacks of traditional eye gaze trackers. However, they are also not well suited for large displays, as head motions are hardly possible and motion within different distances or angles in front of the display is not allowed at all.

A different approach to eye gaze tracking, which we also use in this paper, is gaze estimation based on the orientation of the user's head. A comprehensive survey of head pose estimation techniques is given in [5], where a taxonomy of approaches to head pose estimation is provided. It includes geometric and tracking approaches to head pose estimation, which are similar to our horizontal and vertical gaze estimation described in Section 4. A closely related paper to the present work is [6], where a stereo-based face tracking system for the use with large displays is described. The system operates in real-time and it does not require a per-user initialization process. However, it only supports one user at a time, and does not allow the user to move in front of the display (i.e. the angle and distance to the camera are restricted). Another approach is presented in [7], where the gaze direction is estimated based on stereo tracking of the face and the pupil locations. The system is also quite accurate, and estimates gazes within a range of $\pm 3.5°$; however, it has similar limitations as the previously mentioned work. A commercial gaze tracking system is the Smart Eye system[1], which has been used in [4] for gaze estimation based on head tracking. This approach seems to be quite accurate, but it requires a per-user calibration, was only tested for distances from 1-1.5 m, and does not seem to support more than one user at a time.

4 Real-Time Gaze Tracking

4.1 General Approach

In the following, several assumptions and requirements for our approach will be presented. They provide a common understanding and allow a classification of the system within the broad field of head pose and gaze estimation, respectively. The most basic assumption is that we divide the display into four quadrants and the system's goal is to predict which one is focused by the viewer. The reason for separating the display into quadrants is basically that we wanted to test our approach with respect to (i) looking to the left and to the right as well as (ii) looking up and down, which requires to distinguish a minimum of four display areas. As our approach is based upon an estimation of the head pose only, under the assumption that the eyes are directed straight ahead, a person has to move the whole head towards the quadrant of interest. However, we believe that this is a natural behavior for large displays, as the display area exceeds the field of vision of the user who is standing in its vicinity.

In the remainder of this section, the technical details of our approach will be explained. According to [5], it can be categorized as a hybrid approach, as the horizontal and vertical gaze estimations are accomplished using different

[1] Smart Eye gaze tracker, http://www.smarteye.se

methods, namely a geometric and a tracking method, respectively. Therefore, the process of calculating and predicting the quadrant which the user is looking at is a combination of two independent processes. Each of them produces a result: in the case of horizontal gaze estimation, the result will be either left or right; in the case of vertical gaze estimation, it will be either top or bottom. The simple combination of the results of those two processes yields the outcome of the overall prediction process, i.e. one of the four quadrants. These two processes are illustrated in Fig. 1a and 1b, and will be described in more detail in the next two subsections.

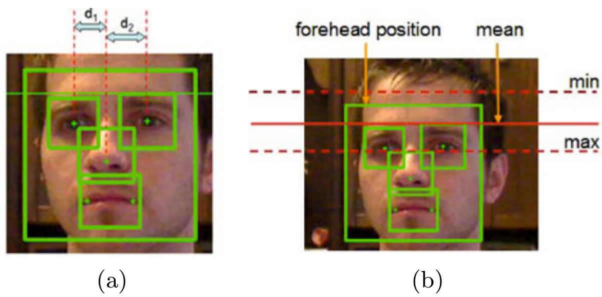


Fig. 1. (a) Horizontal gaze estimation. d_1: distance right eye - nose center. d_2: distance left eye - nose center. *feature-diff* $= d_1 - d_2$. prediction: *left* if *feature-diff* > 0, *right otherwise*. (b) Vertical gaze estimation. Forehead-, minimum- and maximum positions and mean of the latter two.

4.2 Horizontal Gaze Estimation

As stated above, the horizontal gaze estimation is accomplished using a geometric approach [5], i.e. the horizontal viewing direction is derived from the relative positions of different facial features. We investigated which of those features – and in which combination – are best suited to accurately predict the horizontal viewing direction. The facial features are calculated by an object recognition engine which will be explained in Section 4.4. Attempts to involve the position of the mouth were discarded, mainly due to its inaccurate detection in the underlying recognition engine. Therefore, in our further investigation, we focused on the eyes and the nose tip, and finally found out that the best distinction between a view to the left and to the right is by considering the distance of the left and the right eye to the nose tip (see Fig. 1a). We refer to the difference of these two distances as *feature-diff* in the rest of this paper. The achieved correct prediction rate with this approach was about 60-80% for a single image, which was not satisfying. Since the distances between the nose tip and the eyes were the best features we have found, we tried to find a way for improving the prediction rate.

When we analyzed the results, we observed a certain shift to the left or to the right of the *feature-diff* measure depending on the environment (especially the camera direction and position) and the spatial relations between the user and the camera (i.e. the angle and distance between them). Hence, the calculated *feature-diff* has to be adjusted at runtime with a value that reflects this shift. In order to be able to calculate this adjustment, a one-time calibration process is required, which will be described in the following.

Calibration and Adjustment. The runtime-adjustment is based on the assumption that the *feature-diff* values should be zero if the user looks at the center of the display. As mentioned above, this is not the case by default due to environmental conditions and the user's position in front of the display. Therefore, we utilize calibration points at different distances and angles as depicted in Fig. 2b.

The first part of the calibration process is to collect *feature-diff* measurements for all calibration points. The user who performs the calibration has to look at the center of the display, and command the system to store the measured *feature-diff* value for each of these points. The second part of the process concerns the estimation of the user's distance to the display, which is also a crucial aspect for the interaction zones of the proposed pervasive advertising application scenario (see Section 2). As distance information is not available at runtime, but needed in order to predict a good adjustment value, we use a coarse estimation of the user's distance based on the face area. For this purpose, the face area is measured at calibration time at distances from one to four meters in intervals of half a meter as depicted in Fig. 2c. Therewith, we acquire a mapping from concrete distances to face areas at calibration time. The estimation of the user's distance at runtime is then achieved by applying the reverse process, i.e. the current face area is mapped onto a distance estimate using a linear interpolation operation; this process is depicted in Fig. 2d. The adjustment value is then calculated at runtime for each single frame depending on the respective position of the user, where the horizontal position can be determined exactly through the center of the user's face, and the user's distance to the display is estimated as explained above.

Fig. 2b shows the calculation of the adjustment value with input parameters p_u (the user's current position) and the 4 surrounding calibration points. The adjustment value is calculated as the weighted sum of the *feature-diff* values at those 4 points. The weight of one point can be calculated from the factors f_x, which is derived from the user's face position in the image, and f_y, which is derived from the user's distance to the display. With the knowledge of the adjustment value, the prediction of the horizontal gaze can be adapted to improve the correct prediction rate. The new value is *adjusted-feature-diff* $= (d_1 - d_2) -$ *adjust*. To further improve this estimate, a simple smoothing operation is applied as the weighted sum over the last 15 *adjusted-feature-diff* values, where the most recent one gets the highest weight down to the oldest one getting the lowest weight.

It should be noted that the described calibration process has to be performed only once per system setup and not per user or session, which has shown to be sufficient in order to achieve good results (see Section 5). However, further investigations of how specific characteristics of the person who performs the

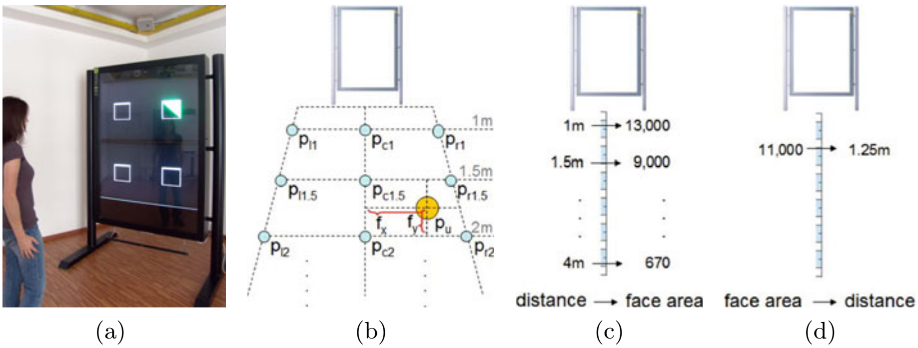


Fig. 2. (a) Prototype setup. (b) Calibration points at different distances (from 1 to 4 m in intervals of 0.5 m) and angles (left, center, right). Position of user p_u and factors f_x, f_y. (c) Mapping, gathered at calibration time, from concrete distances to face areas. (d) Distance estimation at runtime as the reverse operation of (c) with interpolation.

system calibration – e.g. his facial features, gender or height – influence the prediction, will be necessary.

4.3 Vertical Gaze Estimation

The vertical gaze estimation follows a quite simple approach and can be categorized as a tracking method [5]. The basis for the prediction of the vertical gaze is the forehead position. Over a history of 15 frames, the minimum and maximum positions of the forehead are stored. The mean of those two values results in a horizontal line as depicted in Fig. 1b. If the current forehead position is above this line, the vertical gaze prediction will be *top*; if it is below, it will be *bottom*. Whenever a new minimum or maximum position occurs, the horizontal line will be adjusted to the mean of the new values. The initial values for the minimum and maximum positions are obtained as the current forehead position minus or plus a certain threshold, respectively. The applied threshold depends on the face area (i.e. the user's distance to the display), since the it has to be smaller if the face area is smaller, as in this case smaller changes in the forehead position will have a higher impact for the prediction and vice versa. After 15 frames, the history of the minimum and maximum positions is deleted, and the initial values are set as described above. Although this approach may sound too simplistic, it turned out that it works quite well as will be shown in section 5.

4.4 Prototype Implementation

In order to evaluate the recognition accuracy of our gaze tracking approach, we have implemented a fully-functional prototype. It uses the OpenCV library[2] for capturing and processing the input stream from the camera and for visualizing

[2] OpenCV Library v2.1, http://opencv.willowgarage.com/wiki/

various types of feedback to the user, and the SHORE object recognition engine[3] of the Fraunhofer institute for the recognition of facial features like the eyes, nose tip and forehead positions. The implemented prototype is deployed on the hardware setup shown in Fig. 2a, which consists of a large display (1162 x 1364 mm, 1920 x 2240 px) and a video camera (Logitech Quickcam Pro 9000, 960 x 720 px) mounted at a height of 2.1 m on top of the display. For each captured image, a prediction of the current gaze is done for each detected user, whereas the steps for processing one image are: (i) capturing the image, (ii) identifying persons in that image, then (iii) performing facial feature recognition using the SHORE engine and gaze estimation as described in the subsections 4.2 and 4.3 for each person, respectively, and (iv) giving feedback to the user by highlighting the predicted quadrant.

5 Evaluation

With regard to the use of our proposed gaze tracking system for large-scale advertising displays (see Section 2), we conducted a user study which provides empirical results about its recognition accuracy. For an initial functional demonstration, we decided to run the study with only one user at a time standing in front of the display. However, it can be extended to moving people as well as a higher number of people, as the underlying recognition engine supports multi-person tracking in real-time.

5.1 User Study

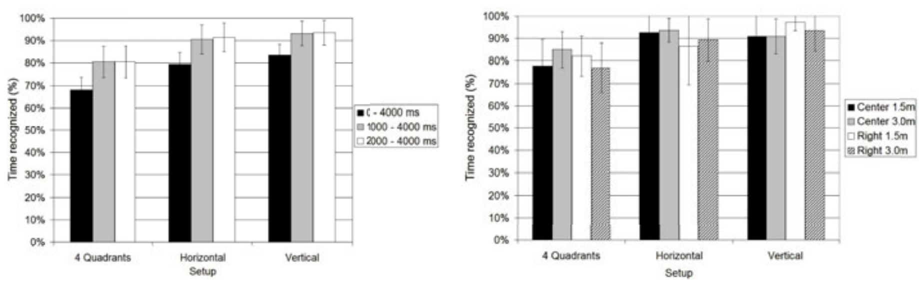
The user study was done with 12 participants, 10 male and 2 female, with a height between 169 and 190 cm. The aim of the study was to explore the accumulated time of correct gaze recognition for three different setups: the recognition of (i) looking to the left versus looking to the right (*horizontal setup*), (ii) looking up versus looking down (*vertical setup*) and (iii) looking at one of four quadrants (*four quadrants setup*). In all three cases, the participants were supposed to look at randomly chosen areas of the display (indicated by a filled triangle as depicted in Fig. 2a) for a certain amount of time. This time period was limited to four seconds per sample, which has shown to be sufficiently long for recognizing the accumulated correct recognition time. For the horizontal and vertical setup the system showed 15 samples per participant, and for the four quadrants setup it showed 30 samples, which is due to the fact that the number of distinguished display areas has also doubled.

Each of the three described setups was executed from different positions in front of the display: 1.5 and 3 m away from the display, as well as in the center and right to the display. With these positions, we cover both the variation in distance and angle of the user in front of the display. Thus, each participant of our study had to conduct 4 x 3 = 12 independent experiments, i.e. the combination of the 3 different setups and the 4 different positions in front of the display.

[3] SHORE, `http://www.iis.fraunhofer.de/bf/bv/kognitiv/biom/dd.jsp`

5.2 Results

The results of the evaluation are shown in Fig. 3a and 3b. The first one shows the mean recognition rate for the three setups and over all 12 participants, and for three different time intervals, namely for starting at 0, 1000 and 2000 ms of the 4000 ms sample period. The reason for the lower recognition accuracy in the former case is the reaction time, which means that it takes some time for the participant to recognize the highlighted area at the display and turn the head towards it. There is no significant difference between the latter two where only the last 3000 or 2000 ms are considered, which supports our limitation of the sample period to four seconds. The lower and higher values of the error bars in the bar chart are one standard deviation each, showing the variance of the recognition times over all participants for a certain experimental setup. The results show (i) that there is no significant difference between the horizontal and vertical recognition accuracy, and (ii) that they are – with about 80% in the four quadrants and more than 90% in the horizontal and vertical setup – well above random results of 25% and 50%, respectively. In Fig. 3b, the results of different spatial relations (distances and angles) of the participants to the display are shown for the 1000 - 4000 ms case. As can be seen in the chart, no significant difference between the four positions can be observed, which supports our initial goal that the recognition accuracy should be invariant with respect to different distances and angles to the display.



(a) Results for different time intervals (b) Results for different positions

Fig. 3. Accumulated correct recognition times per experimental setup

6 Conclusions

In this paper, we presented a real-time gaze tracking approach for public displays, and motivated its importance for estimating the focus of visual attention in pervasive advertising scenarios. It is based on the relative positions of facial features, which are recognized from a single camera mounted on top of the display. In contrast to other systems, it uses head pose estimation only to recognize the user's point of gaze, which has the advantage that it works well for both large distances and angles to the display, and that it is robust with respect to different heights of the users. We have built a prototype system and conducted a user study with 12 participants, which showed that an average accuracy of about 80% can be achieved in the distinction of four different display areas and for different positions in front of the display, and more than 90% if just two display areas are distinguished. The presented work is a promising approach towards gaze-sensitive public displays, where especially an invariant recognition accuracy for different spatial relations between user and display is a core requirement.

References

1. Proc. of the 2nd International Workshop on Pervasive Advertising, in conjunction with Informatik 2009, Lübeck, Germany (2009)
2. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering 17, 734–749 (2005)
3. Morimoto, C.H., Mimica, M.R.M.: Eye gaze tracking techniques for interactive applications. Computer Vision and Image Understanding 98(1), 4–24 (2005)
4. Mubin, O., Lashina, T., Loenen, E.: How not to become a buffoon in front of a shop window: A solution allowing natural head movement for interaction with a public display. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5727, pp. 250–263. Springer, Heidelberg (2009)
5. Murphy-Chutorian, E., Trivedi, M.M.: Head pose estimation in computer vision: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(4), 607–626 (2009)
6. Nakanishi, Y., Fujii, T., Kiatjima, K., Sato, Y., Koike, H.: Vision-based face tracking system for large displays. In: Borriello, G., Holmquist, L.E. (eds.) UbiComp 2002. LNCS, vol. 2498, pp. 152–159. Springer, Heidelberg (2002)
7. Newman, R., Matsumoto, Y., Rougeaux, S., Zelinsky, A.: Real-time stereo tracking for head pose and gaze estimation. In: Proc. of 4th Conference on Automatic Face and Gesture Recognition (FG 2000), pp. 122–128. IEEE CS, Los Alamitos (2000)
8. Ranganathan, A., Campbell, R.H.: Advertising in a pervasive computing environment. In: Proc. of the 2nd Workshop on Mobile Commerce (WMC 2002), pp. 10–14. ACM, New York (2002)
9. Schmidt, A.: Implicit human computer interaction through context. Personal and Ubiquitous Computing 4(2/3) (2000)
10. Strandvall, T.: Eye tracking in human-computer interaction and usability research. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5727, pp. 936–937. Springer, Heidelberg (2009)
11. Vogel, D., Balakrishnan, R.: Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In: Proc. of UIST, pp. 137–146. ACM, New York (2004)

An Agent-Based Approach to Care in Independent Living

Boštjan Kaluža[1], Violeta Mirchevska[2], Erik Dovgan[1], Mitja Luštrek[1], and Matjaž Gams[1]

[1] Jožef Stefan Institute, Department of Intelligent Systems, Jamova cesta 39, 1000 Ljubljana, Slovenia
[2] Result, d.o.o., Bravničarjeva 11, 1000 Ljubljana, Slovenia
{bostjan.kaluza,violeta.mircevska,erik.dovgan}@ijs.si, {mitja.lustrek,matjaz.gams}@ijs.si

Abstract. This paper presents a multi-agent system for the care of elderly people living at home on their own, with the aim to prolong their independence. The system is composed of seven groups of agents providing a reliable, robust and flexible monitoring by sensing the user in the environment, reconstructing the position and posture to create the physical awareness of the user in the environment, reacting to critical situations, calling for help in the case of an emergency, and issuing warnings if unusual behavior is detected. The system has been tested during several on-line demonstrations.

Keywords: Multi-agent system, fall detection, disability detection, independent living, remote care.

1 Introduction and Related Work

The population of developed countries is aging rapidly, increasing the pressure on the working-age population to take care of the elderly. However, given the choice, most elderly people would prefer to continue to live in their own homes (aging in place) [13]. Unfortunately, the majority of elderly gradually lose their ability to function and require either additional assistance in the home or a move to an eldercare (care for the elderly) facility. Moreover, the fear of falling decreases the quality of life and increases the decline in the ability to perform daily activities.

Several systems were introduced in recent years to address some of the issues related to eldercare, e.g., fall detection. However, most of the systems developed in this context are either too expensive for mass use or of low quality. Most commercial solutions are capable only of fall detection, meaning that they recognize only a small set of hazardous situations.

Pan et al. [15], for example, presented a homecare service that tackles the problem of fall discovery using a tri-axial accelerometer and reporting such a discovery to an emergency center. The paper introduces a fall detector based on a neural network and a multi-agent architecture for requesting emergency services. Similar functionalities (fall detection and reporting) were also presented in other papers [2, 5, 6, 9, 18, 20] and systems [1, 19].In addition, Isern et al. [8] reviewed several agent-based solutions for healthcare concluding that such an approach has positive affects in terms of

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 177–186, 2010.

modularity, efficiency, decentralization, flexibility, personalization, distributed planning, monitoring, pro-activity and security. In the remote care field they gave special mention to Agineru [17], a multi-agent system that monitors and processes psychological parameters on a portable device and enables physicians to access the data through a web application based on a centralized database. Besides, Koutkias et al. [10] presented a multi-agent system for the management of chronic diseases that detects anomalous cases and informs personnel; and Cervantes et al. [3] proposed a similar system to [10] for pervasive management that collects and evaluates physiological data collected by sensor agent in order to detect user's symptoms by patient's agent and to report on the situation to doctor's agents.

In the work described in this paper we aim to expand the scope of detection of hazardous situations and augment the detection system by user-adaptation. Our system presents a part of the EU FP7 project Confidence [4, 12], which main objective is to build a care system for the elderly people. An important advantage of our system is that it is based on localization hardware (that enables context-depended reasoning resulting in a lower false-alarm), which is, compared to the video-based motion-tracking systems, low-cost solution and therefore, less accurate. Nevertheless, we developed a low-cost system which accuracy is comparable to high-cost video-based systems. In addition, the localization system is especially suitable for elderly people concerned about their privacy since it is less intrusive than video-based systems. The development of the low-cost less-intrusive system is the main contribution of our work. In addition,unlike most previous work, which basically dealt with just one task, the objective of our research was twofold. The primary goal was to create a care system that is able to detect hazardous situations in short-term behavior, such as falls or a loss of consciousness, which is one of the most common tasks in eldercare. The second, rarely addressed goal was to monitor the user's behavior over longer periods of time and detect changes that indicate decreased performance. Gait disorders, for example, are an important indicator of a variety of health problems: joint and muscular diseases, neurological disorders such as Parkinson's disease etc.

The rest of the paper is organized as follows. Section 2 gives a detailed description of the architecture of our system. Section 3 shows how all the agents work together in two test scenarios: one focused on fall detection and one dealing with detection of long-term changes in behavior. Finally, Section 4 concludes the paper and outlines future work.

2 System Architecture

The system is based on the following six requirements. First, the system is required to monitor a person and detect an emergency situation in real time. Second, the system must be hardware independent, meaning that it can be easily integrated with various localization systems as the sensing component, and coupled with various schemes for user interaction as the communication component. Third, the data presentation must be increasingly more abstract, allowing an inference about the person in the environment at several abstraction levels. For example, a hazardous fall can be detected both from raw sensor data as a sudden change of vertical acceleration, and at the abstract level as a person that laid down quickly in the kitchen. Fourth, the system must allow the introduction of redundancy by combining several methods, sensors and viewpoints to solve given task. This contributes to the stability and robustness of the system as stated by the so-called principle of multiple knowledge. And finally, the system must be able to provide an insight into the person and the environment accessible to an observer.

A multi-agent architecture meets the system requirements. Each system module, task or activity is designed as an agent, providing a service. Agents are organized into groups at a specific level of abstraction and coordinated by another, hierarchically higher-level agent. Each agent can be simply modified or replaced and new or redundant agents can be easily incorporated. An agent provides a service that is requested by another agent or is triggered by some event. The data are available in the agent's output queue waiting to be used, requested or executed by another agent. Agents share their data through direct acquisition or by depositing and searching for the data in the system ontology.

With respect to the goals and requirements we present the architecture illustrated in Figure 1, which organizes the agents into seven groups. Agents from one group have the same architecture and perform logically similar functions, e.g., all sensor agents retrieve the data from sensors. The arrows indicate the communication between agents.

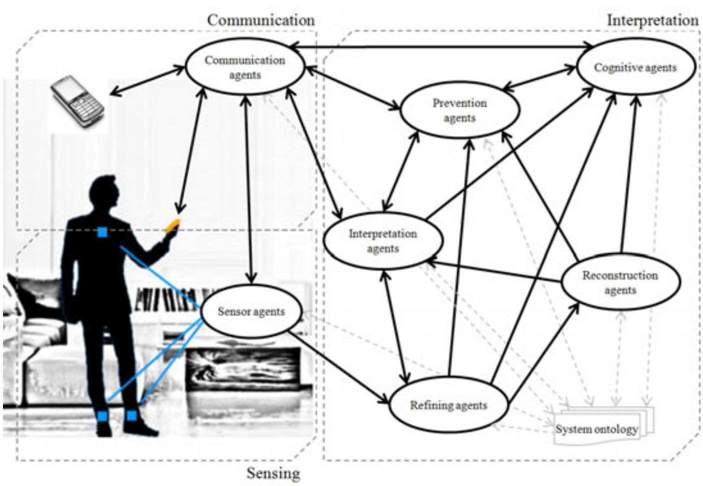


Fig. 1. The agent architecture reveals the various groups of agents. The arrows indicate the communication between agents.

Sensor agents. An arbitrary localization hardware is employed as a sensing agent that serves raw data to the next group of agents. For example, an agent obtains the coordinates of the tag attached to the left ankle. The location agents usually report the sensor data in sequential time periods (1/10 second), but can also report other information such as "low battery".

Refining agents. The refining agents, the second agent group, represent the first abstraction layer. These agents first filter the noise using six independent methods, the most interesting of which is probably a method that corrects coordinates of body parts based on anatomic constraints. The agents then derive attributes describing the user's posture thatform the foundation for the uniform representation of all the available data of the user's body.

Reconstruction agents. The reconstruction agents determine the location and posture of a person at a specific time. These agents provide essential information to activate other agents that assess the situation in the environment and accomplish the actions that are required. There are two groups of prediction agents that work in parallel: the one consisting of machine-learning agents and the one with expert-knowledge agents. The machine-learning agents compute a set of attributes (e.g., distance between tags, velocities, angles etc.) and feed them into Random Forest classifier [11]. The expert-knowledge agents are designed a set of rules proclaiming and negating specific posture [14]. The outputs of both groups are merged by the meta-prediction agents based on heuristics, and learning agents that use, for example, Hidden Markov Models to predict and smooth the final classification.

Interpretation agents. The fourth group, the interpretation group of agents, tries to figure out if a situation is potentially dangerous for the user. For example, it observes that the user is lying immovable at a position where he or she is usually not lying and provides the answer to why this happened – there might be a problem, e.g., the person might have lost consciousness. This group consists of expert knowledge agents, and prediction agents based on machine-learning algorithms, while the final decision is made by a meta-prediction agent. The expert-knowledge agents contain the expert know-how related to the detection of emergency situations that may be caused by a fall or a sudden health problem. These agents are able to recognize four types of emergency situation: (1) falling detected and user lying/sitting immovable at an inappropriate place; (2) user lying/sitting immovable at an inappropriate place for long period of time; (3) falling detected and user lying/sitting at an inappropriate place for a long period of time; and (4) user lying/sitting at an inappropriate place for a very long period of time. The machine-learning agents detect emergency situations based on models induced with machine-learning techniques. Unlike the expert-knowledge agents, these agents predict only the presence or absence of the emergency situation. Their reasoning is using information about the percentage of time the user was involved in specific activities detected by the reconstruction agents, as well as the percentage of time the user was immovable during given time intervals. There are two machine-learning agents, one using SVM machine-learning algorithm and the other using C4.5 algorithm. The final prediction is fused by the integration agent using heuristics.

Prevention agents. The fifth group is the prevention-agent group. Its main task is to prevent development of a disability by monitoring how the person moves at various levels. Agents in this group observe the user's behavior; each of them collects a specific subset of the behavior data in the form of attribute vectors. They automatically build behavior models that are constantly updated using these data.

Consequently, behavior models are sets of the most recent behavior attribute vectors. In particular, the agents observe the dynamics of the person's performance during the carrying out an activity, a task and even the performance during day. Since each agent only partially observes user's behavior, an integration agent collects their observations and merges them into the final behavior observation. The group consists of the following agents: (1) gait-characteristics agents are focused on a person's walk in order to recognize the change in gait using attributes such as the time the user's foot is on or off the ground, step time, step length, etc; (2) turning-characteristics agents are focused on change in direction (e.g., angular velocity, radius of the turn, speed, etc); (3) activity-characteristics agents observe the general behavior during various activities using a set of simple attributes, for example, the general speed during walking and lying, the walking speed, the time of posture transitions during sitting down and standing up, etc; and (4) spatial-activity agents are focused on the daily behavior of a person thus observing a person over a longer period of time, i.e., the dynamics of a person in the apartment by monitoring activities performed during the day combined with spatial information. Each of the prevention agents sends the messages asynchronously to the integration agent, which collects the messages for a predefined period. Afterwards, it analyses the outlier degrees of all the observed behaviors, i.e., the degrees to which they deviate from the usual behavior. If an outlier degree is higher than the high warning threshold, a warning is sent to the user.

Cognitive agents. The next group is the cognitive-state agent group: an increasingly more abstract context awareness makes it possible to construct the cognitive state of the integrated system. The system design includes the cognitive state of the user (although not implemented yet in the tested version). The cognitive layer will also include the attributes, related to the cognitive state of the user, thus constructing not only the physical, but also the cognitive state of the user, using a similar agent architecture as with the lower levels. However, these agents will use the cognitive state to perform reasoning on a wider spectrum of information with an integrated reasoning strategy.

Communication agents. The last group contains communication agents dedicated to user interaction, for example, agents that alert the person with a reply demand, make a phone call to relatives or help center, graphically display the state of the system etc. An example of the communication agent is an alarm agent that triggers when a dangerous situation is reported by other agents and shuts down when the user responds that everything is OK. We have developed a set of graphical and communication agents: (1) an agent to inform the person about an alarm or warning; (2) an agent that interacts with the user (to raise, cancel or hold an alarm); (3) graphical agents that explain the current state of a particular agent group; (4) agents that are able to explain why a specific alarm or warning was raised. Currently, our prototype implementation runs entirely on a PC. We are in the process of extending it with advanced services and devices such as an emergency call service, a portable device for alarm inhibition, a mobile phone, a remote web-based monitoring application etc.

3 Experimental Results

For the prototype deployment we have equipped a room as a near-realistic apartment of about 25 square meters. The test room contained a bed, a few chairs and tables, and was divided into six logical areas: a kitchen, where the user can prepare a meal; a sleeping area, devoted to sleeping; a living room, where the person can eat a meal, watch TV, write a letter etc.; a toilet; and a corridor, an area where the user enters and leaves the room.

Since the equipment with which the Confidence system will acquire locations is still under development, the commercially available localization system Ubisense [16] was used instead. Ubisense allows local positioning by tracking a set of reasonably small tags, which are attached to a person. A sampling frequency of around 10 Hz can be achieved with no more than four tags attached to a person simultaneously. Each tag maintains radio contact with a sensor mounted on the wall. These sensors use ultra-wideband (UWB) technology to detect the position, and make use of both the time difference of arrival and the angle of arrival to calculate the location. In a typical open environment, a location accuracy of about 15 cm can be achieved across 95% of the readings [16], although it can occasionally significantly drop. Due to the fact that the Ubisense sampling rate of 10 Hz is limited to up to four tags, we decided to position the tags at the following locations on the body: chest, belt (optional), left and right ankle.

We have designed two sets of experiments showing the capabilities of the interpretation (fall detection) and prevention (disability detection) group of agents. Both experiments are described in the following subsections.

3.1 Fall-Detection Experiments

The first experiment explores the fall-detection capabilities of the interpretation group of agents. For this purpose, we designed a complex scenario including three situations when an alarm must be raised and three situations where a false alarm is likely. The situations in which the alarm must be raised are tripping (quick fall), fainting (slow fall, potentially not detected by accelerometers) and sliding from the chair (sliding slowly from the chair and sitting on the ground, again potentially not detected by accelerometers). The situations that might lead to a false alarm, but are actually safe, non-alarming situations, are jumping in a bed, sitting down quickly and searching under the table/bed. Both jumping in a bed and sitting down quickly contain falling, thus potentially misleading the system. For systems based on accelerometers this is the primary problem.

These two sets of situations basically reveal the ability of the interpretation-agent group to carry out context-dependent reasoning, i.e., to distinguish between an alarm situation and normal behavior consisting of the same user activities based on the context in which they are done. For example, searching under the table/bed may be falsely perceived as prolonged lying on the ground, since the person is close to the ground. However, it differs from lying on the ground by the amount of user movement and the length of time the user is on the ground.

The prototype was tested on recordings of five people, each performing the above-mentioned test scenario five times, which is 25 recordings in total. The evaluation was performed with the leave-one-person-out methodology.

Table 1 presents the achieved accuracy of the agents in the reconstruction agent group. The first three rows represent the alarm situation (A), where the alarm must be raised, the next three rows represent false-alarm situation (FA), where alarm should not be raised, and the last row represents the overall accuracy. The columns represent the accuracy achieved by machine-learning and expert-knowledge agents, while the last column represents the results achieved by the integration agent.

Tripping and fainting were recognized quite reliably in both cases – the alarm was always triggered. What proved problematic was detecting slipping from a chair. It consists of sliding down slowly from the chair and sitting on the ground. Due to the noise in the data from the localization system, which is particularly high near the ground, sitting on the ground and sitting on a chair is sometimes difficult to distinguish. Therefore, the system is not sure what is actually the right posture or even believes that the person is moving, due to the noise.

Table 1. Performance of the reconstruction agent group when detecting alarm and non-alarm situations.

Case \ Accuracy [%]	Machine-learning agents	Expert-knowledge agents	Meta-prediction agents
A: Tripping	40.00	92.00	100.00
A: Fainting standing	64.00	100.00	100.00
A: Sliding from the chair	32.00	52.00	52.00
FA: Jumping in bed	100.00	96.00	100.00
FA: Sitting down quickly	100.00	100.00	100.00
FA: Searching for something under a table/bed	96.00	88.00	96.00
Overall	**72.00**	**88.00**	**91.33**

The system raised only one false alarm. Difficulties were observed when searching under the bed/table. The main cause for this is the inability of the activity-recognition agents to accurately distinguish a person on all fours from a person lying on the ground.

In summary, the system performed quite well. The problems it had were rather exotic, e.g., falling from a chair and remaining sitting or being on all fours for a certain period, which is also a rather unusual situation, in particular for the elderly.

3.2 Disability Detection Experiments

The second experiment tested the ability of the prevention group of agents to detect changes in behavior, indicating present or emerging disease or illness. The testing of the general disability/disease detection would ideally require many days of recordings of daily activities. Such tests are currently in progress, but initially we condensed a full five days of activities into recordings that lasted around 25 minutes each. Two days of normal activities serving as a training period were recorded first, followed by a day of normal and two days of abnormal activities served as a testing period.

The agents in the prevention-agent group monitor a variety of statistics of daily living and aggregate them over specific periods. If the period over which the statistics are aggregated is shortened, we expect the condensed-day tests to be a reasonable substitution for multiple-days tests. The scenarios for the first two days represent a typical daily routine for an elderly person. There is some variation in activities, but not much in general terms since the elderly tend to follow a routine. On the fourth day the elderly person is not feeling well and as a consequence is moving slowly and rests a lot. Such behavior could occur if he/she had flu or any other general health problem. On the fifth day the elderly person is limping. As a consequence, the elderly is also moving slowly and is avoiding having to stand. The person is not lying as much as on the previous day, but is sitting more than normal.

The results of the tests are shown in Table 2. The days were divided into five-minute observation periods (five periods each day). The periods in which each group of agents raised a warning are shaded. Despite the small number of days used in the training phase, the initial results are very promising: the prevention mechanism perceived normal behavior (white fields in Table 2) on the normal day, while on the slow and limping days, each prevention agent raised a warning at least once, often more than once and in most periods more than one agent raised a warning (gray fields in Table 2). It should be noted that the simulations were not performed by an elderly person, but by healthy persons supervised by a professional physician. The simulations were as real-life as possible for the distinctive changes in performance. This is potentially the reason for the 100% performance of the system.

Table 2. Performance of the prevention-agent group during a normal day, limping and slower movement due to a health problem. The detections are marked with gray.

<table>
<tr><th>Day part / agents</th><th>Gait characteristics agents</th><th>Turning characteristics agents</th><th>Activity characteristics agents</th><th>Spatial-activity agent</th><th>Meta-prediction agent</th></tr>
<tr><td>Normal 1</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Normal 2</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Normal 3</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Normal 4</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Normal 5</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Slow 1</td><td style="background:gray"></td><td style="background:gray"></td><td></td><td></td><td style="background:gray"></td></tr>
<tr><td>Slow 2</td><td></td><td style="background:gray"></td><td style="background:gray"></td><td></td><td style="background:gray"></td></tr>
<tr><td>Slow 3</td><td></td><td style="background:gray"></td><td style="background:gray"></td><td></td><td style="background:gray"></td></tr>
<tr><td>Slow 4</td><td></td><td style="background:gray"></td><td></td><td></td><td></td></tr>
<tr><td>Slow 5</td><td></td><td style="background:gray"></td><td style="background:gray"></td><td style="background:gray"></td><td style="background:gray"></td></tr>
<tr><td>Limping 1</td><td></td><td style="background:gray"></td><td style="background:gray"></td><td></td><td style="background:gray"></td></tr>
<tr><td>Limping 2</td><td></td><td></td><td style="background:gray"></td><td></td><td></td></tr>
<tr><td>Limping 3</td><td></td><td></td><td style="background:gray"></td><td></td><td style="background:gray"></td></tr>
<tr><td>Limping 4</td><td></td><td style="background:gray"></td><td style="background:gray"></td><td></td><td style="background:gray"></td></tr>
<tr><td>Limping 5</td><td style="background:gray"></td><td style="background:gray"></td><td></td><td style="background:gray"></td><td style="background:gray"></td></tr>
</table>

4 Conclusions and Future Work

In this paper we have presented a multi-agent system that consists of seven groups of intelligent agents, i.e., sensor, refinement, reconstruction, interpretation, prevention, and cognitive group of agents. The agents within groups are arranged horizontally, contributing various interpretations of the situation, and vertically, providing increasingly more abstract situational awareness. Each agent in group has its own strong and weak points, while the advanced combination and integration overcomes the particular weaknesses and combines different aspects into a reliable interpretation.

The results in the fall-detection experiment show that context-dependent reasoning can interpret complex scenarios that might be misinterpreted by acceleration-based systems. In addition, the preliminary results on the disability detection are encouraging, showing a potential for the early discovery of a potential health problem that may lead to perilous conditions. Overall, hundreds of hours of tests were already performed, including an on-line presentation to EU reviewers with 100% performance both in the prepared scenarios and in the scenarios defined on-the-spot. However, the real applicability of the presented system will be demonstrated in the final year of the project, when it will be installed and tested in the homes of elderly people.

In order to increase the usability, performance and effectiveness of the multi-agent system, our work in the immediate future will focus on the following goals. The first goal is an initialization procedure that will optimize the initial parameters of the system to a particular user, for example, the user person will show how he/she usually performs some activities, specify his/her height in order to refine the agents in the reconstruction mechanism, set parameters that will describe his/her normal-day dynamics in order to relax or strengthen the alert sensitivity, set privacy-related services for reporting, monitoring, etc. Secondly, we will develop additional procedures for user adaptation during the system operation, in particular, dynamic adaptation to false alarms and undiscovered alarms. Finally, we will devise effective and pro-active agents to broaden the scope of the awareness and prevention mechanism. Ultimately, the success of the interpretative mechanism's ambitious goals depends strongly on the ability to suit and adapt to an end user.

Acknowledgments. This work was partly supported by the Slovenian Research Agency under the Research Programme P2-0209 Artificial Intelligence and Intelligent Systems, and partly from the European Community's Framework Programme FP7/2007–2013 under grant agreement No. 214986. The work of Violeta Mirchevska was partly financed by the European Union, European Social Found. We would like to thank Domen Marinčič, Rok Piltaver, Boža Cvetkovič, Blaž Strle and Damjan Kužnar for suggestions, discussions and help with the programming.

References

1. AlertOne Services Inc, iLife™ Fall Detection Sensor, `http://www.falldetection.com`
2. Bourke, A.K., Lyons, G.M.: A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor. Medical Engineering & Physics 30(1), 84–90 (2008)

3. Cervantes, L., Lee, Y.-S., Yang, H., Ko, S.-H., Lee, J.: Agent-based intelligent decision support for the home healthcare environment. In: Szczuka, M.S., Howard, D., Ślęzak, D., Kim, H.-k., Kim, T.-h., Ko, I.-s., Lee, G., Sloot, P.M.A. (eds.) ICHIT 2006. LNCS (LNAI), vol. 4413, pp. 414–424. Springer, Heidelberg (2007)
4. Confidence, `http://www.confidence-eu.org`
5. Doughty, K., Lewis, R., McIntosh, A.: The design of a practical and reliable fall detector for community and institutional telecare. Journal of Telemedicine and Telecare 6, 150–154 (2000)
6. Fu, Z., Culurciello, E., Lichtsteiner, P., Delbruck, T.: Fall detection using an address-event temporal contrast vision sensor. In: Proceedengs of the IEEE International Symposium on Circuits and Systems, pp. 424–427 (2008)
7. Fiol-Roig, G., Arellano, D., Perales, F.J.L., Bassa, P., Zanlongo, M.: The intelligent butler: A virtual agent for disabled and elderly people assistance. In: DCAI. Advances in Soft Computing, vol. 50, pp. 375–384. Springer, Heidelberg (2008)
8. Isern, D., Sanchez, D., Moreno, A.: Agents applied in health care: A review. International Journal of Medical Informatics 79(3), 145–166 (2010)
9. Kangas, M., Konttila, A., Winblad, I., Jamsa, T.: Determination of simple thresholds for accelerometry-based parameters for fall detection. In: Proceedings of the 29th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, pp. 1367–1370 (2007)
10. Koutkias, V., Chouvarda, I., Maglaveras, N.: A multiagent system enhancing home-care health services for chronic disease management. IEEE Transactions on Information Technology in Biomedicine 9(4), 528–537 (2005)
11. Lustrek, M., Kaluza, B.: Fall detection and activity recognition with machine learning. Informatica 33(2), 197–204 (2009)
12. Lustrek, M., Kaluza, B., Dovgan, E., Pogorelc, E., Gams, M.: Behavior analysis based on coordinates of body tags. In: Tscheligi, M., de Ruyter, B., Markopoulus, P., Wichert, R., Mirlacher, T., Meschterjakov, A., Reitberger, W. (eds.) AmI 2009. LNCS, vol. 5859, pp. 14–23. Springer, Heidelberg (2009)
13. Means, R., Richards, S., Smith, R.: Community Care: Policy and Practice. Palgrave MacMillan, Basingstoke (2008)
14. Mirchevska, V., Lustrek, M., Gams, M.: Combining machine learning and expert knowledge for classifying human posture. In: Proceedings of the Eighteenth International Electrotechnical and Computer Science Conference, vol. B, pp. 183–186 (2009)
15. Pan, J.-I., Yung, C.-Y., Liang, C.-C., Lai, L.-F.: An intelligent homecare emergency service system for elder falling. In: World Congress on Medical Physics and Biomedical Engineering, pp. 424–428. Springer, Heidelberg (2006)
16. Steggles, P., Gschwind, S.: Ubisense – a smart space platform. Technical report, Ubisense (2005)
17. Tablado, A., Illarramendi, A., Bagues, M.I., Bermudez, J., Goni, A.: Aingeru: An innovating system for tele-assistance of elderly people. Journal on Information Technology in Healthcare 2(3), 205–214 (2004)
18. Willis, D.J.: Ambulation Monitoring and Fall Detection System using Dynamic Belief Networks. PhD Thesis, School of Computer Science and Software Engineering, Monash University (2000)
19. Zenio: Zenio Fall Detector, `http://www.zenio.be/product/8.html`
20. Zhang, T., Wang, J., Liu, J., Hou, P.: Fall detection by wearable-sensor and one-class SVM algorithm. In: Lecture Notes in Control and Information Science, pp. 858–863 (2006)

Making AAL Platforms a Reality

Antonio Kung and Bruno Jean-Bart

Trialog, 25 rue du Général Foy, 75008, Paris, France
antonio.kung@trialog.com, bruno.jean-bart@trialog.com

Abstract. This paper takes the position that the advent of Ambient Assisted Living (AAL) depends on the availability of an ecosystem based on widely accepted mainstream platforms supporting AAL domain specific interfaces. It describes the features which such platforms must meet, based on the experience of the MonAMI project. It identifies challenges concerning the standardisation of application specific interfaces, and their support by platforms. It finally suggests two measures that will help make such platforms a reality: first, the launching of a research and industry initiative for an open source platform; secondly, the launching of a EC policy supported long-term process for the adoption of AAL standards.

Keywords: AmI, AAL, Platform, Applications, Policies, Standards.

1 Introduction

The term Ambient Intelligence (AmI) refers to "a vision of people surrounded by intelligent intuitive interfaces that are embedded in all kinds of objects and an environment that is capable of recognizing and responding to the presence of different individuals in an invisible way" [1]. AmI systems are complex and sophisticated. They involve applications which must be embedded, context-aware, personalized, adaptive, and anticipatory. They necessitate underlying technologies such as unobtrusive hardware, seamless communication, dynamic distributed networks, human-centric computer interfaces and dependable trusted systems.

Ambient Assisted Living (AAL) refers to "intelligent systems that will assist elderly individuals for a better, healthier and safer life in the preferred living environment and covers concepts, products and services that interlink and improve new technologies and the social environment" [2]. AAL systems address socio-political issues created by the demographic development in Europe. Care and assistance needs to elderly persons will drastically increase in the future with a situation in 2050 where there will be two retired persons for one working person. The 2006 Riga ministerial declaration on e-Inclusion policy [3] has recognised AAL systems as a social necessity but also an economic opportunity. This has led to the definition by the European Commission of an ageing well action plan [4] as well as a series of measures that involve more than one billion euro in research and development between 2006 and 2013.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 187–196, 2010.

AAL systems are ambient intelligent systems. They share the same vision and involve the same key technologies. Their advent will therefore benefit from progress made on such systems, but also from the availability of specific features and standards. This paper takes the position that an ecosystem based on widely accepted mainstream platforms while supporting AAL domain specific interfaces is needed. To make such platforms a reality, we suggest two measures, first, the launching of a research and industry initiative for an open source platform and secondly, the launching of a EC policy supported long-term process for the definition of AAL solutions and standards.

The rest of the paper is structured as follows. The next section provides an overview of AAL platforms. We then list challenges that need to be solved concerning them. We finally make suggestions on a roadmap for AAL and on two supporting measures. This paper is an extension of [5].

2 AAL Platforms

AAL platform can be defined as a computing architecture including software and hardware that serves as a foundation to application programmers. For instance, iOS [6] is the platform used for the development of iPhone applications. Platforms allow for separation of concern between features implemented by a small community of system experts and application features implemented by a larger community of developers, thereby allowing much faster development of applications. Platforms must evolve and mature as the applications they support significantly evolve, as witnessed by the many evolutions of Windows for the development of PC applications.

In order to transform the today fragmented markets into a unified market, we need platforms for ambient intelligent applications that will provide the needed support for hardware, communication, distributed networks, user interfaces, dependability to allow for the effective development of applications that are context-aware, personalized, and anticipatory. And we need to leverage on future AmI platforms while specialising them into AAL platforms that can be adopted by the community of AAL applications developers.

2.1 MonAMI Contribution to Platforms

MonAMI [7] is an acronym for Mainstreaming on Ambient Intelligence, and as its name suggests, the overall objective of this FP6 project is to investigate, specify, validate and promote an approach based on mainstream technologies whereby AAL applications can be deployed in a cost effective manner.

In order to understand how an overall AAL system could be suitably structured, MonAMI undertook a deconstruction exercise. Starting from the layered-based TAHI open architecture [8], MonAMI identified four building blocks, the content management, the service management, the premises infrastructure, and the network management building blocks (see figure below). Content management provides *quality of content*. It involves the creation, collection, protection of data which are of interest to an application. It could be sensor data, or video data. Service management provides *quality of service*. It exploits content in order to provide application features. It includes context management and therefore integrates the core AAL applications. The premises infrastructure provides *quality of access*. It groups devices and equipment in the home. Finally network management refers to the underlying communication infrastructure. It provides *quality of network*.

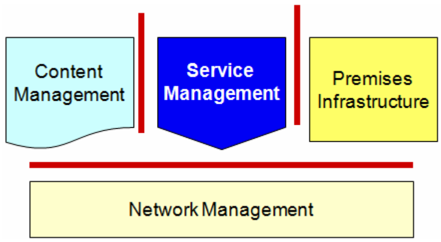


Fig. 1. AAL applications deconstructed into four blocks

Further to this deconstruction exercise, MonAMI undertook a platform analysis process with the objective to favor features that would make the business environment for developing AAL services simpler, which would in turn foster the creation of a business ecosystem. By business environment, we refer to the community of stakeholders involved in the development and operations of AAL services, including decision makers.

The first feature selected was the use of well defined interfaces between applications and the platform, so that independence from the overall technology infrastructure is ensured. In practice this would mean that application developers would not have to adapt their code to specific devices and technologies.

The second feature was the definition of an appropriate architecture. Three architectures were initially identified, the *web-centric*, *home-centric* and *fully distributed* configurations. While each configuration includes web level computing elements and home-based computing elements, they have differences in terms of weight.

Web-centric configurations consist of "fat" servers directly connected to home sensors and actuators. In such configurations, end-user applications mainly run in the remote server. We believe that such configurations could create barriers. Web-services programming requires know-how that is generally not easily accessible by a large community of developers. Further, data concerning end-users are collected and aggregated at the web-server level which involves serious privacy issues.

Home-centric configurations consist of "thin" web servers and "fat" home servers. Web servers focus on administration aspects mainly, such as subscriber's management. End-user applications are executed by dedicated devices, or home servers which directly access to home sensors and actuators. Home-centric configurations are generally based on PC centric platforms and therefore more easily accessible by the community of developers. Furthermore, deployment of applications is a well known process, e.g. in PCs where is software routinely updated. Finally processing of user data stays local and therefore limits privacy issues.

In fully distributed configurations, there is no notion of fat or thin nodes, applications run anywhere, in the home and elsewhere. Despite the fact that such configurations are well understood and technically mature, there is no widely available mainstream technology, except possibly for peer-to-peer systems.

Consequently, MonAMI decided to focus on the home-centric architecture, as it creates a much simpler environment for developing services compared to approaches where a service is deployed at the web server level, or even worse when the service consist of pieces that need to be deployed both at the web server level or at the home device level.

MonAMI finally developed a proof of concept platform. The following technologies were selected: a service framework based on OSGi [9], an accessibility framework based on URC (Universal Remote Console) [10,11], and sensors and actuators based on ZigBee [12]. OSGi is a Java-based service platform that can be flexibly managed remotely. URC provides for flexibility of access. ZigBee is a radio-frequency (RF) solution based on the IEEE 802.15.4 standard for wireless personal area networks.

A set of interfaces called OSGi4AMI [13] was developed to create the clear separation between applications and the set of devices (e.g. sensors, actuators) used in MonAMI. OSGi4MI involves two levels of interfaces, the device level and the function level (see Figure below).

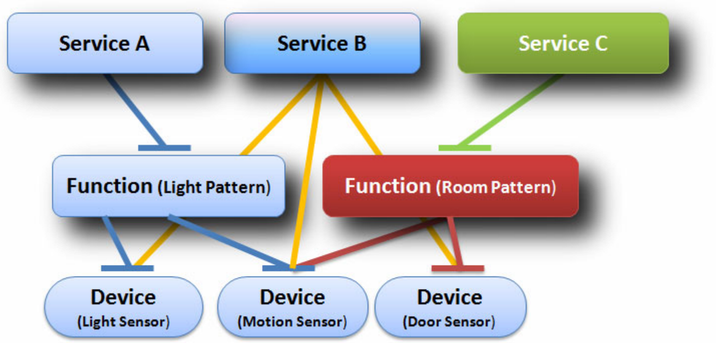


Fig. 2. MonAMI two-level interface

An initial implementation of the platform was carried out in six locations (France, Germany, Slovakia, Spain, Sweden, UK) with the implementation of six applications: AMiSURE (Monitoring and alerts), AMiCASA (Controlling home environment), AMiVUE (Remote home surveillance), AMiPAL (Reminder system) and AMiPLAY (Entertainment and physical stimulation). These implementations not only showed the usability and effectiveness of specified applications, but they also confirmed that a clear interface between applications and the underlying technology platform was crucial for rapid and straightforward implementation of applications. MonAMI is still an on-going project as it will now deploy the implemented applications in a living scale field trial involving 50 homes in Slovakia, Spain and Sweden in order to assess real condition usability and scalability aspects.

2.2 Other Platforms

Several initiatives specific to AAL have been launched to address the issue of creating open service platforms. SOPRANO [14] uses an OSGi-based home platform as well as a sensor/actuators ontology and a context ontology. PERSONA [15,16] also uses OSGi technology. The service/platform interface is also based on ontology. It defines four interfaces, the input bus, the output bus, the service bus and the context bus. MPower [17] is based on SOA (service oriented architecture). OASIS [18] investigates the use of ontology at all levels and is OSGi-based. I2Home [11] has developed an open accessibility framework based on the URC specification [10]. Finally UniversAAL [19] is a recently started project which intends to consolidate current results and promote open platforms.

2.3 The Service-Platform Approach

MonAMI and the initiatives mentioned above have all focused on the service/platform interface. This single interface, if properly defined could facilitate the transformation of the today vertical markets in a single unified horizontal market. It would also facilitate user centred design as application development iterations could be made possible by allowing faster implementation and modifications of applications. Finally it would ensure that applications would be technology independent. For instance OSGI4AMI is independent from ZigBee, so other networking technologies could be reused without changing the implementation of applications.

We are only half-way through. The resulting platform developed by MonAMI is open, i.e. based on existing open mainstream technology, and similarly OSGi4AMI is an open specification. This is the same case in the other development initiatives. But this is not sufficient to kick start the ecosystem for AAL applications. On the one hand, an agreement on which home platform to use could just be out of the scope of the AAL community. It would rather be the result of pushes from mainstream applications and technologies. OSGi is a very probable candidate but other platforms could be available in the future. Many programming platforms of today for nomadic devices, or even console game could just be as relevant. On the other hand, we need to reach an agreement concerning the interface between services and platforms. And this agreement is AAL specific. OSGi4AMI is a good starting point, but it is dedicated to the six types of applications developed in the project, therefore it is not comprehensive. Further it has not been validated nor agreed by stakeholders in the application area who are not involved in MonAMI. To our knowledge there is no discussion on the definition of a common overarching interface which would enforce a clear separation between stakeholders developing applications and stakeholders developing platform elements. We believe this is a prerequisite for an AAL ecosystem.

3 Challenges for AAL Platforms

This section lists five challenges that we have to cope with concerning AAL platforms: the integration of innovation in platforms, the integration of transversal features, interoperability of interfaces, technology independence and finally the support of multiple business models.

3.1 Integrating Innovation

While AAL platforms should be based on widely used technologies, they do include features that are specific and novel. The AAliance research agenda [6] provides a good overview of all the challenges that are at stake in different application areas (at home, within a community, at work). It identifies technology challenges related to sensing, reasoning, acting, interacting, and communicating. Continuous research today and tomorrow will lead to further progress in those area. The impact on existing platforms should be assessed. The challenge is to maintain and coordinate two threads of development activities, on one hand the integration of research features into a platform so that they can be quickly tested and validated, and on the other hand the integration of components in an platform that is industrially deployed.

This is not the case today. Open source implementations today such as OpenAAL [20] are research platforms which are not meant for use in industry deployment.

3.2 Integrating Transversal Features

Future AAL platforms should take into account features that are transversal. Scalability issues will arise when millions of platforms are deployed and personalised. Quality of service issues will arise for specific applications with different dependability requirements, e.g. related to health monitoring or alarms. Liability issues will arise when the chain of applications and technology has not worked properly. Confidentiality and data protection issues may also arise that will have far reaching effects on application and platform design.

3.3 Interoperability

Interoperability is not easy to reach. First of all there are many domains. The Interoperability working group of the German BMBF/VDE Innovation Partnership on AAL [21] includes for experts from the following domains: Home/building automation, medical IT, household appliances, consumer electronics, network technology, AAL middleware platforms, sensor networks, robotics, cognitive systems, standardisation and certification.

Secondly, any attempt for standardisation based on consensus building will involve one or even several dedicated task forces *per domain* involving the various competing stakeholders in the domain. Experience shows that in order to get consensus the agenda of such working group should involve discussion on competitive issues, i.e. what is mandatory, what is optional, what is proprietary, and on how new extensions can be provided to an interoperability standard. This was the case of the CHAIN standard [22], a group which involved more than 10 white goods manufacturers over a 3-year plus undertaking in order to build a consensus on interoperability at the semantic level, as well as on the evolution of related ontology [23].

The AALiance project has also identified areas where there is no standard: “Remote Maintenance of AAL Systems, Connecting to Medical Sensors, AAL

Terminology, vocabularies, and ontology's, AAL Middleware / Service Execution Environment, Emergency Calls and Connection to Call Centres" [24,25].

We identify two scenarios for successful standard adoption: in the first scenario, a proposed standard is based on individual initiatives. Success depends on the very individual stakeholders which have created the standard. This is the case of the the Continua Health Alliance [26] which while not based on overall consensus, does include a significant list of organisations with strong commitments. Note that Continua does mention "aging independently" as one of the three domain that it will support. The second scenario consists in reaching overall consensus between all stakeholders within an application domain and technology stakeholders responsible for a given platform. We believe that this second scenario is what is needed for AAL platforms.

3.4 Technology Independence

One can wonder why technology independence is a needed requirement. After all, at some point one has to agree on the use of a platform and commits to the set of technologies that are associated. In the case such platform involves physical interactions, then interoperability can only take place if (1) the application / platform is standardised, and (2) involved devices support the underlying communication protocol used.

But technology independence serves other purposes. First of all, it allows applications stakeholders to switch from one platform to another if there are several competing platforms. Secondly it allows platform stakeholders to switch from one technology to another (for instance to another RF network solution).

The challenge is to express the interfaces in a technology independent manner. This was the approach used in the CHAIN specification [22] which made sure that interoperability specifications would consist of three parts: a network technology independent specification of semantic interoperability (e.g. the "start appliance" message is technology independent), a network technology independent specification of syntaxic interoperability (e.g. the "start appliance" message is represented by value 1), and finally a network dependent mapping (e.g. the message is mapped on a Konnex network). A working group in CENELEC has just issued a document on IFRS, an interoperability framework requirements specification which provides the "language" for expressing semantic interoperability independently of the underlying network standard [27]. Such framework could be used in the future for technology independent specifications.

3.5 Multiple Business Models

There could be several business models depending on national/regional specificities. In some countries, AAL applications could be subsidised at the public level (e.g. social security), and in other countries this could be at the private level (insurance company). Local policies have an impact on the way AAL applications and platforms are procured and operated. The resulting different business models could have an impact on the service/platform interface requirements. For instance different technology features would be needed depending on whether sharing of data between different domains (e.g. care versus health) is possible or not. The challenge therefore is to ensure that the specified standards and interfaces be compatible with the various operating models.

4 Elements of a Roadmap and Related Measures

This section describes a roadmap and two related measures that will contribute to making AAL platforms a reality.

4.1 MonAMI Roadmap

MonAMI has defined a roadmap consisting of 3 phases. The first phase is *a proof of concept phase*. It relates to the work carried out in MonAMI and in other R&D projects. They have led or will lead to proof of concept solutions in terms of applications, platforms, interface specifications, based on assumed business models. The second phase is *a transformation phase*. It must involve initiatives for the creation of AAL platform communities which federate as much as possible the various contributions in Europe. They must be accompanied by policy measures that ensure that a consensus building community is set up so that suitable interfaces between applications and platforms are defined. In parallel, individual ventures can take place. The combination of initiatives for AAL platforms communities, of measures for consensus building and of individual ventures will lead to the third phase, *the ecosystem phase*. A vibrant ecosystem is available with strong business undertakings.

4.2 Need for a Platform Community

The first identified initiative is the set up of a platform community which federates the input of all the research projects while preparing the conditions so that the platform can be industrially be used. This is not an easy undertaking because conflict of interests could arise between research and industry priorities. Therefore the governance to the two types of activities must be carefully discussed. AALOA [28] and OpenURC [29] are two open source initiatives which are currently under creation, and we hope that proper discussion for governance will take place taking into account both research and industry viewpoints.

4.3 Need for a Consensus Building Community

The second identified initiative is the creation of an overarching consensus building community with a long term perspective. The objective would be to help define technology independent specifications as well as their evolutions as applications and technology for AAL evolve, over a time span of several decades. This measure could be implemented under the responsibility of the European Commission as part of the first action of the ageing well action plan set [4] which states the following: “The Commission will therefore facilitate efforts of business stakeholders and civil society organisations to establish an innovation platform for ageing well”.

The Commission is familiar with this type of measure. For instance, article 16 (2) of the IPPC directive on pollution prevention [30] states that the “Commission shall organise an exchange of information between Member States and the industries concerned on best available techniques, associated monitoring, and developments in them”. This has been achieved through the IPPC bureau in Sevilla which consists of a staff of 20 persons belonging to the competitiveness and sustainability unit. The bureau is in charge of organising a consensus building process (called the Sevilla process [31]). This process involves currently 30 working groups.

Finally we would like to point out that the cost of such measures is small compared to the budgets that have been allocated to R&D projects.

5 Conclusion

MonAMI has tried to solve two problems at the same time. The first problem was to identify meaningful AAL services, assuming that a suitable platform was available. The second problem was to identify meaningful platform features, assuming that AAL services were available. This difficult undertaking has allowed us to get an understanding of the measures that will be needed to make AAL platform a reality.

We would like to acknowledge the work and contribution of MonAMI partners, as well as the positive feedback obtained through the various interactions with the AAL community and the European Commission. MonAMI is partially funded by the European Commission under the framework 6 program.

References

All URLs have been viewed on June 2010.

1. Information Society Technology Advisory Group (ISTAG) Scenarios for Ambient Intelligence (2010), ftp://ftp.cordis.lu/pub/ist/docs/istagscenarios2010.pdf
2. Ambient Assisted Living Roadmap. AAliance, http://www.aaliance.eu
3. Ministerial Declaration, Riga, Latvia, June 11 (2006), http://ec.europa.eu/information_society/events/ict_riga_2006/index_en.htm
4. Action Plan on Information and Communication Technologies and Ageing, June 14 (2005), http://ec.europa.eu/information_society/activities/einclusion/policy/ageing/action_plan/index_en.htm
5. Kung, A., Fagerberg, G.: Alliance for an AAL Open Service Platform. In: AALIANCE Conference - Malaga, Spain, March 11-12 (2010), http://www.aaliance.eu
6. iPhone OS Reference Library, http://developer.apple.com/iphone/library/navigation/
7. MonAMI IST project, http://www.monami.info
8. TAHI: The Application Home Initiative, http://www.theapplicationhome.com/NewFiles/TAHI-TOA-Lite-V7.pdf, http://www.theapplicationhome.com
9. OSGi, http://www.osgi.org
10. Universal Remote Console, http://myurc.org/

11. I2Home IST project, http://www.i2home.org
12. http://www.zigbee.org
13. Marco, A., Casas, R., Bauer, G., Blasco, R., Asensio, A., Jean-Bart, B.: Common OSGI Interface for Ambient Assisted Living Scenarios. Ambient Intelligence and Smart Environments 3 (2009)
14. SOPRANO IST project, http://www.soprano-ip.org
15. PERSONA IST project, http://www.aal-persona.org
16. Furfari, F., Tazari, M.-R.: Realizing Ambient Assisted Living Spaces with the PERSONA Platform. ERCIM News 74 (July 2008)
17. MPower IST project, http://www.sintef.no/Projectweb/MPOWER
18. OASIS IST project, http://www.oasis-project.eu
19. UniversAAL IST project, http://www.universaal.org/
20. Open AAL open source middleware, http://openaal.org/
21. BMBF/VDE Interoperability working group, http://www.vde.com/de/Technik/AAL/Publikationen/Kongress-undFachbeitraege/Seiten/BuchInteroperabilit%C3%A4tvonAAL-Systemkomponenten.aspx
22. CECED Chain specification, http://www.ceced.org/
23. Kung, A.: Architecture and Design Patterns for Ambient Intelligent Systems: an Industry Perspective. In: Proceeding of the Amid 2007. Springer, Heidelberg (September 2007)
24. Ambient Assisted Living Strategic Research Agenda. AAliance, http://www.aaliance.eu
25. Policy Paper on Standardisation Requirements for AAL. AAliance, http://www.aaliance.eu
26. http://www.continua.org
27. Interoperability framework requirements specification for service to the home (IFRS). Cenelec Workshop Agreement CWA 50560 (June 2010)
28. AALOA, http://www.aaloa.org
29. OpenURC, http://www.openurc.org
30. Directive 2008/1/EC concerning integrated pollution prevention and control, January 15 (2008)
31. Schoenberger, H., European Commission: Integrated pollution prevention and control in large industrial installations on the basis of best available techniques – The Sevilla Process. Journal of Cleaner Production 17, 1526–1529 (2009)

A Unified Architecture for Supporting Direct Tag-Based and Indirect Network-Based Resource Discovery

Simone Mora[1] and Babak A. Farshchian[2]

[1] Norwegian University of Science and Technology
[2] SINTEF ICT
Trondheim, Norway
Simone.Mora@idi.ntnu.no,
Babak.Farshchian@sintef.no

Abstract. Discovering and integrating ambient computational resources is a central topic in AmI. There are two major existing approaches: indirect network-based resource selection and direct tag-based resource identification. We motivate the need to integrate the two approaches through a scenario. We then present an architecture for a pluggable discovery system called UbiDisco. We demonstrate how UbiDisco implements a seamless integration of the two approaches at user interaction level through a framework for implementing discovery actions.

1 Introduction

AmI environments often contain embedded resources that can be used to enhance the experience of the user, or to improve the performance of the underlying system in one way or another. In order to allow dynamic and unobtrusive access to resources in the environment, applications and services need to reconfigure continuously based on the current context of the users as the users move around. As the number of network-connected resources grows it becomes quite time-consuming for users and system administrators to manage meaningful connections among resources. The trade-off between the effort to find and connect to the right resources in one hand, and the perceived value gained from these resources on the other hand, is a central research question for us.

Automation of resource associations in an AmI system requires a first essential step, i.e. discovery and integration of the resource into the AmI system. Dynamic resource discovery systems have therefore shown to be a major cornerstone of all AmI and ubicomp systems [3]. Conventional resource discovery architectures have been used to connect computers. They are therefore implemented with network topology in mind. AmI systems require a higher degree of user-centeredness. It is the intention of the user that is in focus when discovering new resources, and not the network topology. A recent line of research has therefore focused on supporting direct user-initiated integration of resources into AmI systems using embedded tags [4][5].

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 197–206, 2010.

A user is requested to read or touch some form of tag (barcode, RFID, IR) embedded into a desired resource. The tag is used to obtain a handle to a digital representation of the resource. Each of the two approaches has its own advantages and disadvantages:

- **Direct tag-based discovery:** The user can directly interpret his/her intention into something that the AmI system can immediately recognize. There is no guessing about what resources the user needs. Moreover, observing resources in the ambient can initiate serendipitous and spontaneous interaction with resources [6]. On the other hand there is little automation in the discovery process. Interaction design becomes a major issue. Another disadvantage is that discovering resources that are not in the line of sight (e.g. a printer in the neighboring room) is not possible.
- **Indirect network-based discovery:** The main advantage is that resources can be discovered and integrated regardless of their location. The disadvantage, or merely the challenge, is to know what resources the user really needs. The AmI system has to guess what resources, among all that are registered, are most relevant to the user in a given, highly dynamic context.

We believe both of these approaches are important to an AmI system, and that the underlying discovery mechanisms should equally support both. But the two approaches are often developed in parallel with little integration. In this paper we will present a system called UbiDisco (Ubiquitous Discovery system) that integrates the two approaches in an open and extensible architecture. In Section 2 we will introduce a scenario and show how an example application based on UbiDisco solves some of the challenges of discovery in AmI. In Section 3 we introduce the architecture of UbiDisco, and in Section 4 we will describe the current implementation of UbiDisco and in Section 5 how it is evaluated so far. Section 6 represents some related work. Section 6 concludes the paper with a discussion of the results so far.

2 UbiDisco Scenario and Main Concepts

This section gives an overview of the main concepts of UbiDisco. UbiDisco is part of the UbiCollab platform [7]. UbiCollab is an experimental platform that aims to support mobile users in remote collaboration. UbiCollab implements the concept of a *human grid*, i.e. a collection of physically distributed people who share a context for collaboration. Participants in a human grid use locally available resources in order to facilitate remote collaboration. Resources play a central role in UbiCollab, and UbiDisco is the subsystem of UbiCollab in charge of discovering ambient and web resources. *Resources* in UbiDisco can be internet-enabled devices, web services, or physical objects with a digital representation. Once a resource is discovered using UbiDisco and integrated into the UbiCollab platform, the resource provides a set of *Services* to UbiCollab *Applications*. Resources are integrated into UbiCollab using *Proxy Services* which are similar to device drivers.

Discovery and integration of resources is done using a set of UbiDisco *Discovery Actions*. A Discovery Action is responsible for guiding the user through a set of interactions in order to obtain a handle to a desired resource. A Discovery Action can be implemented to acquire the handle directly from the resource (using e.g. a tag) or to look up the user's query in an online catalog. The Discovery Action framework of UbiDisco enables a seamless combination of direct tag-based discovery and indirect network-based discovery. The interaction mechanisms implemented by UbiDisco provide a seamless integration of the two also at the user interface level.

2.1 Scenario

This section describes a typical scenario for UbiDisco.

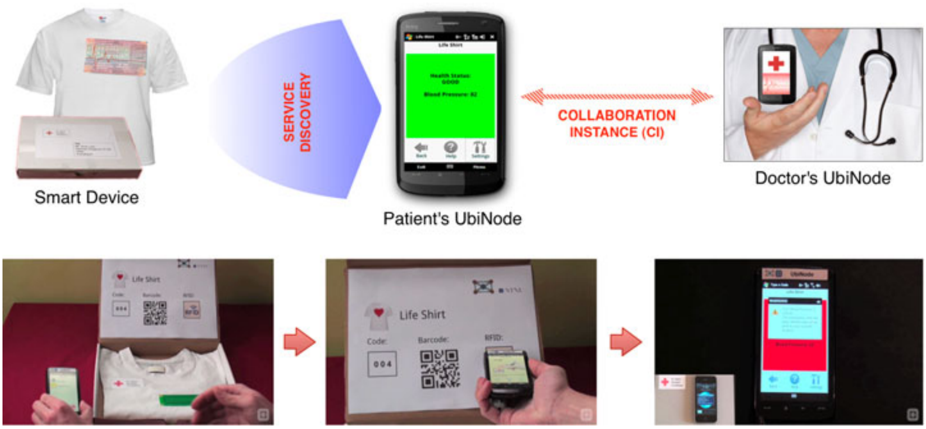


Fig. 1. Scenario illustrating the discovery of lifeshirt by reading an RFID tag

Arne has a chronic heart condition. He is equipped with a UbiNode from his hospital. It is a smart touch-based device that he uses in order to keep in touch with the hospital. The UbiNode has a number of applications installed. One is where he registers his physical condition for his doctor to monitor. During his latest visit to the hospital he is prescribed a lifeshirt, a new intelligent shirt with embedded sensors. After unpacking the shirt, he finds a sheet of paper in the box with instructions about how to use the shirt. It is written that he should add the shirt to his UbiNode. He can choose among three ways of adding the device: 1) take a photo of a barcode on the paper, 2) touch an RFID tag on the paper, 3) punch in a number written on the paper. He clicks on "Add Service" on his UbiNode. Then he selects "Touch RFID" action, and points his UbiNode to the sheet of paper. His UbiNode presents a number of screens with some information and a confirmation page. Then a message shows up telling him that the device was added successfully. The shirt now communicates with his UbiNode using Bluetooth. Arne then starts his medical application. The application notifies him that a lifeshirt is detected and asks him whether he wants to use it.

Some days later, Arne is on a regular visit to his doctor. In the hospital he decides to print the latest measurements he has done. He cannot find a printer service on his UbiNode. He clicks "Add Service" on his UbiNode because he knows he can use it for finding printers. He inspects the possible actions. All of them require him to have a printer close to him. Since he cannot see a printer around, he cannot use any of the actions on his UbiNode. He looks around and finds a tag on the wall with a label "Read this tag to access hospital services." He knows that the tag will give him new options for discovery actions. He clicks "Add new action" on his UbiNode, and takes a picture of the tag on the wall. After confirming on his screen, he now finds a new action in the list of Add Device actions that reads "Search for Hospital's services". He selects the action, and can browse in a list of services, one of them reading "Print documents". He selects the option and is presented with a printer name, and a link to a map that shows where he and the printer are. He follows the map, finds the printer, and adds it to his UbiNode. Next time he knows where to print his documents.

2.2 Analysis

The main point of the above scenario is to demonstrate that resource discovery in AmI needs to be user-centered, i.e. taking users' situation and needs as the starting point. Service discovery is not an act that can be fully automated. Intelligence should be introduced where there is a clear need (demonstrated by the system finding which printer is closest to Arne). In particular the scenario demonstrates why discovery of resources in AmI should be supported both using direct tag-based and indirect network-based discovery mechanisms. Direct tag- based discovery utilizes the properties of the physical space such as proximity and visibility[8], and eliminates a number of risks for "misfiring" i.e. finding the wrong resources.

Although the above points are well-known issues, there are a number of additional innovations represented by our approach. Using a framework for Discovery Actions, and a uniform user interaction mechanism, UbiDisco abstracts away the differences between finding a resource using direct tag-based and indirect network-based discovery. The user is represented by a list of "Actions" that includes examples of both approaches. The goal is to "Add Service". Some actions involve physical activity (e.g. taking a photo or touching an RFID tag) while others involve a mental activity (e.g. browsing a list of resources). The user goes through similar and consistent interaction mechanisms.

3 UbiDisco Architecture

UbiDisco (being part of UbiCollab) employs a peer-to-peer philosophy where all users are equal peers in a human grid, and are represented by a personal device called UbiNode (see Fig. 2). UbiDisco denotes the combination of Resource Discovery Manager and discovery plugins (shown as pieces of puzzle in Fig. 2).

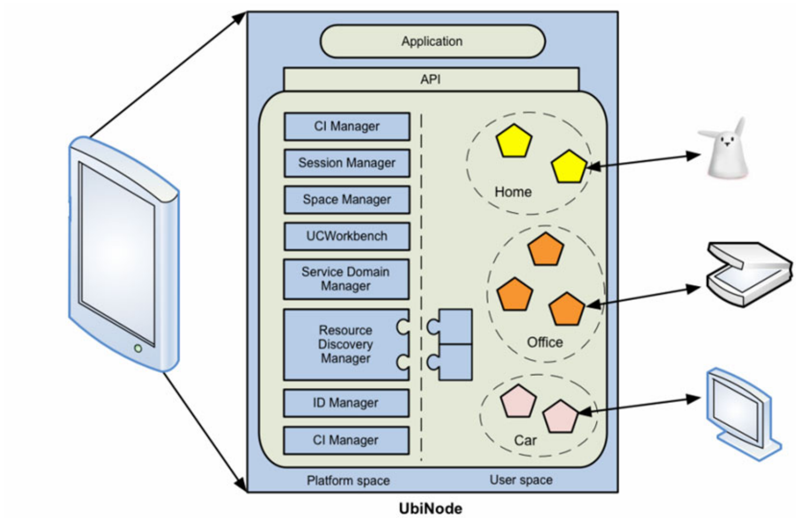


Fig. 2. The architecture of UbiNode

3.1 Overall Service Management Framework

UbiDisco assumes the existence of a proxy layer towards external resources. For each resource that the user wants to use, UbiDisco requires the availability of a Proxy Service. A Proxy Service is code that acts more or less as a traditional driver. Once a Proxy Service is installed on user's UbiNode, UbiCollab applications can use the associated resource. The following steps are performed by UbiCollab service management framework (see Fig.3 on next page):

- **Resource discovery:** Implemented by UbiDisco. During this step a user (or an application) identifies the need for a specific resource. The need is specified in form of a query, and is provided to UbiDisco. A set of discovery plug-ins are notified and start searching for resources. The search can involve various Discovery Actions implemented by each of the plug-ins (Step 1 in Fig.3). Each resource that is discovered is stored in a resource pool along with the ID of the application that asked for the resource.
- **Proxy Service installation:** Implemented by UbiCollab's Service Domain Manager (SDM). The URI obtained during the resource discovery phase by UbiDisco points to an XML document that describes the resource and provides a link to download a Proxy Service for the resource. The URI is passed to SDM (Step 2 in Fig.3). SDM downloads the Proxy Service (Step 3 in Fig.3) into user's Service Domain on the UbiNode (Step 4 in Fig.3). Once the Proxy Service is installed it provides a service interface that can be accessed by applications (Step 5 in Fig.3).
- **Sharing of services:** As an optional step, the user can choose to share a Service Proxy with other users. In this way these users can get access to the local resource. Sharing of resources is a central part of the UbiCollab platform, but is out of scope for this paper and will not be discussed further. Please refer to [7].

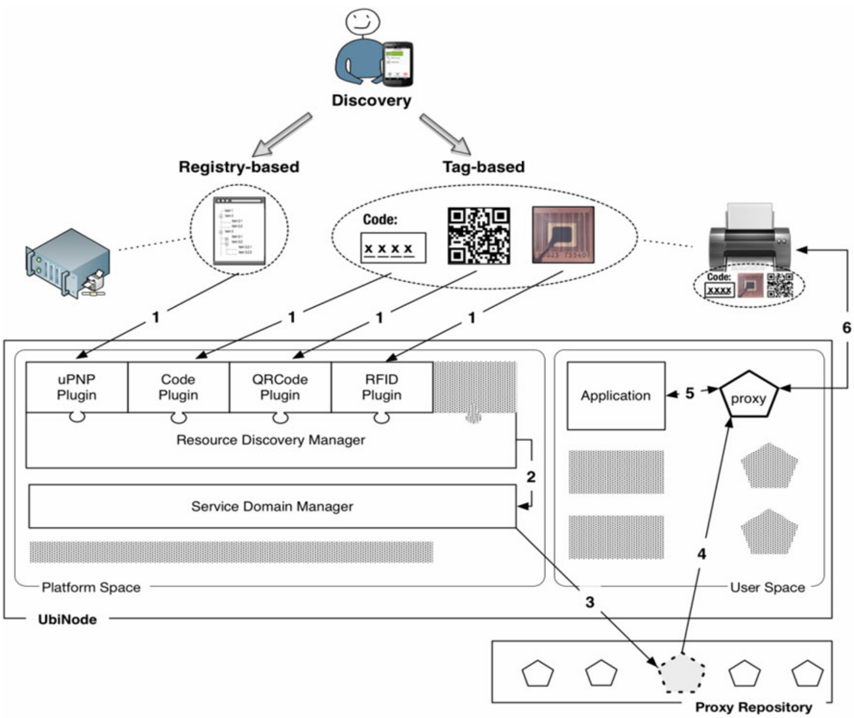


Fig. 3. Overall discovery process in UbiDisco

3.2 Discovery Plug-ins and Discovery Actions

As shown in Fig. 3, the resource discovery phase of UbiDisco is implemented by a set of what we call discovery plug-ins. Plug-ins are developed by third parties and adhere to the framework defined by the UbiDisco. This framework mainly requires an API to allow clients send requests for resources to be discovered, and the implementation of a Discovery Action user interface. Each plug-in that is installed on user's UbiNode will attempt to find the resource according to its own Discovery Action and discovery protocols. The returned value is a set of handles to discovered resources. A plug-in module consists of two parts: a network communication part (implementing the specific discovery protocol such as RFID-based or UPnP), and a Discovery Action part.

At the core of UbiDisco are the various Discovery Actions (DAs) that illustrate innovative ways of user-centered discovery in AmI. A DA is the physical or mental act the user needs to perform in order to be able to start using a newly encountered resource. Various types of DAs can be observed in many existing AmI systems but a coherent framework is lacking. In UbiDisco, the implementation of a DA is done by the discovery plug-ins in a uniform way controlled by SDM. This provides a platform for experimenting with various actions and discovery mechanisms. The following actions have been currently developed for UbiDisco:

- **Touching RFID tags:** The user starts an application that needs a printer. The user observes a printer with an RFID tag. The RFID action is selected by the user, and the user touches the tag with his/her UbiNode. The printer is installed and used by the application.
- **Type a number:** Same as the above, but the action shows a keyboard asking the user to type the number of the printer. The user observes a number tag on the printer and types in the number.
- **Take photo of 2D barcode:** Same as above, but this time the user uses UbiNode's camera to take a photo of the barcode attached to the printer.
- **UPnP:** The user is represented with a list of UPnP services in the network and is asked to select the exact resource.

4 Implementation

UbiCollab platform components, including UbiDisco, are all components that share the same internal architecture. These components are encapsulated into OSGi bundles[1]. User components (applications, proxy service and discovery plug-ins) are also OSGi bundles that are installed via a Discovery Action, and run without rebooting the JVM due to dynamic deployment features of OSGi.

A UbiCollab component contains two layers: the *Platform Abstraction Layer (PAL)* and the *User Abstraction Layer (UAL)*. The PAL contains the basic logic and connects the component to other components and to the OSGi framework. The UAL connects the module to a *User Interface Manager (UIM)*, which in the current implementation is represented by the UCWorkbench; a module built on a eWorkbench draft provided by the Eclipse Foundation. In discovery plugins, PAL implements network communication while UAL implements the discovery action.

UIM is the central user interaction module in UbiDisco (and in UbiCollab). It is the glue between user actions and the underlying component model (see Fig.4). It provides a plug mechanism that allows third-parties to develop Discovery Actions without extensive coding.

A *View* in UbiCollab architecture implements a unit of interaction corresponding to a user task (lowest layer in Fig.4). Similar Views can be grouped into *Perspectives*, which roughly correspond to user intentions (middle layer in Fig.4). For instance, if user's intention is to discover a resource, the user will switch to the Discovery Perspective of UbiCollab, implemented by UbiDisco. There are a number of perspectives in UbiCollab that correspond to different perceived user intentions.

Discovery plug-ins implementing the Discovery Actions have been designed exploiting well-known low-cost object tagging standards as RFID and QRCode tags. The RFID plug-in makes use of high-frequency tags (13.56 MHz), suitable to be scanned by portable readers. Resources URIs are stored in a registry embedded in each discovery plug-in, to allow *off-line* discovery.

---

[1] OSGi is a Java-based container technology. See www.osgi.org.

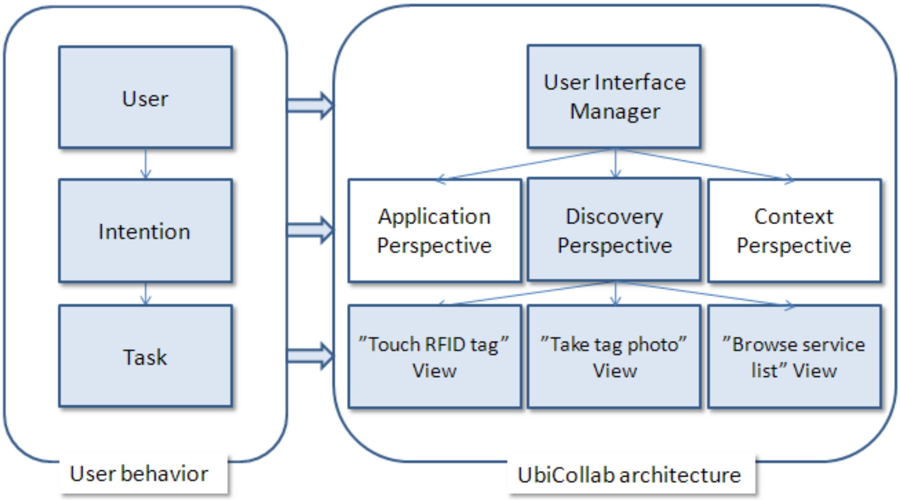


Fig. 4. User Interface Manager manages perspectives and views related to user intention

5 Preliminary Evaluation Results

We are in the early stages of evaluating our results on real users. UbiDisco and the implemented discovery actions have been subject to focus group and hands-on testing by end users of primarily technical background. The groups have varied from 5 to 14, and feedback is collected using Technology Acceptance Model. Preliminary results show potential for improvements in the usability area. The results from these studies have indicated that the relationship between applications and resources is an area that needs more research. The transition from usage perspective to discovery perspective has shown to be confusing to our users. Discovery actions are however considered useful, and the most popular one among our users has so far been the RFID touch action. A more scientific analysis of user feedback will be the subject of a future paper.

6 Related Work

Discovering and integrating resources with users' applications as users move around introduces a number of challenges that are addressed by researchers in the recent years. There are a number of classification schemes documented in the literature that demonstrate a great variety in how resources are represented, how registries are constructed, and how look-up criteria are decided in indirect network-based discovery [3][9]. From being purely network-centred, indirect network-based discovery has become increasingly user-centred. Examples include selection of services based on user activity[10], and selection based on user preferences[11] or context elements such as location and spatial relations[12]. There exist also a number of commercial systems that implement community-based service registries, such as AppStore, Android Market and various UDDI-based directories.

Direct tag-based discovery was pioneered among others by the work on CoolTown[4] and has been developed into its own field of research with an increasing number of interesting results. Augmentation of everyday objects with discovery mechanisms is demonstrated in [5][13]. In [14] an interesting approach is demonstrated where multiple tag-based discovery actions are combined to implement associations among devices. Our research builds on this existing research and extends it with pluggable architectures that allow integration of the two approaches.

7 Conclusions and Future Work

In AmI environments saturated with computational resources, discovery and integration of resources becomes a major undertaking for the users. In professional environments, such as offices, the knowledge at least exists among the users about how to configure and connect resources to the AmI system (but the time to do so is often lacking). In other scenarios, such as ambient assisted living, we cannot assume enough knowledge or capabilities exist among the elderly. In all cases, there is a great potential to automate or simplify a lot of the routine work involved in finding and connecting to the right resources. We have demonstrated how various discovery actions can be integrated into UbiDisco architecture in order to adapt discovery to capabilities and preferences of users and make resource discovery more user-centered.

We are in the process of experimenting with more innovative discovery actions, such as voice-based and the usage of pointing devices. Security and privacy issues haven't yet been considered, but our future plans in this area include digital signing of proxies and plug-ins, and the usage of virtual identities.

References

1. Streitz, N.A., Rocker, C., Prante, T., van Alphen, D., Stenzel, R., Magerkurth, C.: Designing smart artifacts for smart environments. Computer 38, 41–49 (2005)
2. Satyanarayanan, M.: Pervasive computing: vision and challenges. IEEE Personal Communications 8, 10–17 (2001)
3. Edwards, W.: Discovery Systems in Ubiquitous Computing. IEEE Pervasive Computing 5, 70–77 (2006)
4. Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B., Spasojevic, M.: People, Places, Things: Web Presence for the Real World. Mobile Networks and Applications 7, 365–376 (2002)
5. Riekki, J., Salminen, T., Alakarppa, I.: Requesting Pervasive Services by Touching RFID Tags. IEEE Pervasive Computing 5, 40–46 (2006)
6. Newman, M.W., Sedivy, J.Z., Neuwirth, C.M., Edwards, W.K., Hong, J.I., Izadi, S., Marcelo, K., Smith, T.F.: Designing for serendipity: Supporting end-user configuration of ubiquitous computing environments. In: Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, pp. 147–156. ACM, London (2002)

7. Farshchian, B.A., Divitini, M.: Collaboration Support for Mobile Users in Ubiquitous Environments. In: Handbook of Ambient Intelligence and Smart Environments, pp. 173–199 (2010)
8. Harrison, S., Dourish, P.: Re-place-ing space: the roles of place and space in collaborative systems. In: Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work, pp. 67–76. ACM, Boston (1996)
9. Zhu, F., Mutka, M.W., Ni, L.M.: Service discovery in pervasive computing environments. IEEE Pervasive Computing 4, 81–90 (2005)
10. Bucur, D., Bardram, J.: Resource Discovery in Activity-Based Sensor Networks. Mobile Networks and Applications 12, 129–142 (2007)
11. Park, K., Yoon, U.H., Kim, S.: Personalized Service Discovery in Ubiquitous Computing Environments. IEEE Pervasive Computing 8, 58–65 (2009)
12. Gellersen, H., Fischer, C., Guinard, D., Gostner, R., Kortuem, G., Kray, C., Rukzio, E., Streng, S.: Supporting device discovery and spontaneous interaction with spatial references. Personal and Ubiquitous Computing 13, 255–264 (2009)
13. O'Neill, E., Thompson, P., Garzonis, S., Warr, A.: Reach Out and Touch: Using NFC and 2D Barcodes for Service Discovery and Interaction with Mobile Devices. Pervasive Computing, 19–36 (2007)
14. Broll, G., Haarländer, M., Paolucci, M., Wagner, M., Rukzio, E., Schmidt, A.: Collect&Drop: A Technique for Multi-Tag Interaction with Real World Objects and Information. Ambient Intelligence, 175–191 (2008)
15. Berger, S., McFaddin, S., Binding, C., Hoertnagl, C., Ranganathan, A.: Towards pluggable discovery frameworks for mobile and pervasive applications. In: Proceedings of IEEE International Conference on Mobile Data Management, 2004, Berkeley, CA, USA, pp. 308–319 (2004)
16. ASTRA, `http://www.astra-project.net/`

Multilevel and Hybrid Architecture for Device Abstraction and Context Information Management in Smart Home Environments

Víctor Peláez, Roberto González, Luis Ángel San Martín, Antonio Campos, and Vanesa Lobato

Fundación CTIC, C/ Ada Byron, 39, Edificio Centros Tecnológicos, Parque Científico y Tecnológico Gijón, Asturias, 33203, Spain
{victor.pelaez,roberto.gonzalez,luisangel.sanmartin}@fundacionctic.org, {antonio.campos,vanesa.lobato}@fundacionctic.org

Abstract. Hardware device management, and context information acquisition and abstraction are key factors to develop the ambient intelligent paradigm in smart homes. This work presents an architecture that addresses these two problems and provides a usable framework to develop applications easily. In contrast to other proposals, this work addresses performance issues specifically. Results show that the execution performance of the developed prototype is suitable for deployment in a real environment. In addition, the modular design of the system allows the user to develop applications using different techniques and different levels of abstraction.

Keywords: smart home, middleware, frameworks.

1 Introduction

In recent years the concept of digital home and the evolution of this concept to the smart home [1] by applying the principles of the Ambient Intelligent paradigm [2] have received a great deal of attention. However, in spite of several research projects developed over the last five years, the impact of this research in homes has been limited [3]. The divergence between research project results and technologies deployed in real houses is the result of several factors:

- Most systems are designed to work with hardware devices not available to the general public (e.g. hardware developed in research projects)
- It is difficult to deploy research results in real environments as few research projects take into account performance or scalability analysis.
- Specific problems of the digital home are not addressed because results from other ubiquitous computing researches are adapted.
- Research outcomes are not freely available to the scientific community.

At this time there is no architecture that has been accepted globally and that covers all these problems (hardware integration, versatility, reusability and performance). In order to solve these problems, this work proposes an architecture (and the associated implementation) with the following advantages:

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 207–216, 2010.

- The architecture supports devices available in the market and addresses the interoperability problem between different families of devices.
- Device abstraction and context information abstraction is supported with different levels of detail
- The system supports the integration of intelligent services (context-awarenes, proactive and adaptable services) based on AI techniques.
- The architecture provides a clear and simple API in order to facilitate the development of applications and services for the smart home using different levels of abstraction and techniques.
- The architecture is able to isolate inefficient parts that could be disabled dynamically. Performance tests have shown that the system can be deployed in real environments

2 Challenges in the Current Digital Home

Nowadays the concept of smart home is understood as the combination of electronic devices, management middleware and user interfaces installed in homes in order to provide different services to users.

One of the main problems to be resolved is the lack of interoperability between different types of devices in the digital home. There are standards for home automation networks (KNX, LonWorks, X10, Dali, etc.), to represent and model sensor devices (oBiX, SensorML, etc.) [4], for multimedia devices (UPnP, DLNA, etc.) or data networks (802.11, Bluetooth, Zigbee, etc.). Therefore the development of services or applications that use all these types of devices together is very difficult due to an increasing number of standards. This interoperability problem is incremented by the fact that digital devices in existing homes are acquired in an accidental and unplanified way in function of the users' needs and without considering technical factors [5].

The main research challenge is the necessity for a general architecture that facilitates the development of intelligent applications independant of the underlying hardware infrastructure. The main problem found in existing state of the art works is the difficult deployment of these solutions in real environments. Some architectures are not validated with real hardware devices [6] [7] or the performance is not analyzed [8] [6]. When it is analyzed, the system has a bottleneck entity [9] [7].

3 Proposed Architecture

3.1 Design Considerations

In home automation networks such as X-10, KNX or LonWorks, devices are connected to a communication network with a bus topology. Each device is able to communicate with other devices by writing and reading messages to and from the bus. In a similar way, more complex devices such as UPnP devices or DLNA devices have similar behavior but interchange messages in XML format over an

IP network. As different types of devices in the home can be controlled by reading and writing bus messages, each message in a physical bus can be abstracted by a message in a logical bus. In addition, each physical device can be represented as a logical service connected to the logical bus.

Even when complex discovery mechanisms or configuration files generated automatically are available, some information about the devices or the environment cannot be obtained automatically, context information must be written manually. For example, it is not possible to determine the location of each device automatically.

3.2 General Design

A multi-level architecture is proposed to separate different levels of abstraction and complexity (see Figure 1). The lower level manages information coming from hardware devices, while the superior level uses a semantic model that simplifies application development. This structure enables the isolation of complex parts and avoids global bottlenecks in the system (e.g. a central repository of information based on a semantic model).

The architecture is divided in three different levels according to the abstraction level of the information used at this level:

- Physical level. The information at this level is represented using the physical formats utilized by the different types of devices.
- Syntactic level. The information at this level is represented in a common format for all the devices of the same type regardless of their manufacturer or model. This level ensures inter-operability between different types of devices by using a common XML vocabulary (described later).
- Semantic level. The information at this level is obtained by enriching the information in the syntactic level with meta-information. This enrichment process gives meaning to the information coming from hardware devices by taking into account the global model of the home. The semantic information is represented with RDF according to an ontology that can be customized or completely user-defined. The semantic information represents not only devices but context information in general.

Information is enriched and abstracted while it is sent from one level to the next one using the different buses of the system. The entities in the system (from physical to semantic level) are described as follows:

- Device Driver. This entity is connected to a specific network of devices and enables the interchange of messages between the *physical bus* and the *physical abstraction bus*. There is a specific driver for each family of devices to be controlled.
- Device Manager. This entity creates a *device service* instance for each hardware device to be abstracted. To create the service instances, the device manager uses the appropriate discovery mechanism or if no discovery techniques are available, the *Devices configuration*. There is a *Device Manager* for each family of devices.

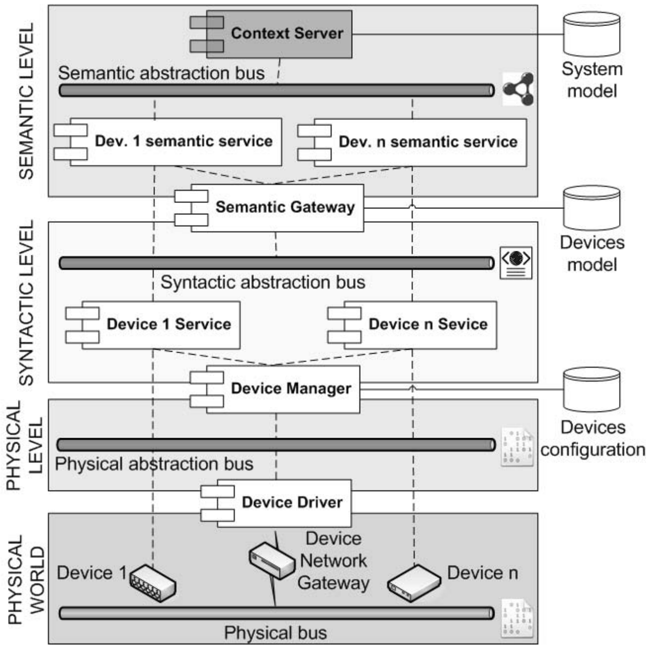


Fig. 1. General architecture

- Device Services. These entities are service instances that create a logical representation for each different device in the physical world. These services read messages from the *physical abstraction bus* to synchronize their state with the state of the physical device. In a similar way services write messages to this bus to execute actions over the physical devices. Devices with the same functionalities are abstracted by services with the same interface regardless of the device model or manufacturer. When the state of the associated device changes, the service sends an event to the *Syntactic abstraction bus*. This event contains a message in a common XML vocabulary that abstracts the hardware device.
- Semantic Gateway. This entity creates a *device semantic service* for each *device service* by using the meta-information available in the *system model*. The system model contains meta-information that must be generated manually as it is related to the context of the devices.
- Device Semantic Service. These services add meta-information, such as the location in the home of the abstracted device, to *device services*. These services also associate semantics to the information coming from hardware devices. This semantic association is done by using the context information available in the *system model*. Finally, these services also convert XML-based events from the *Syntactic abstraction bus* into RDF-based events with the associated meta-information. These RDF-based events are sent to the *semantic abstraction bus*.

– Context server. This entity receives all the events sent to the *semantic abstraction bus* by *device semantic services* and aggregates all the RDF information in a single dynamic model. The context server also has a static model of the system (information about rooms, users, etc.) represented in RDF following a specific configurable ontology. Thus it creates a centralized model that represents the current state of the system (i.e. context information). As the context server works as a event listener, it can be disabled easily.

It is important to note that the context server is used to convert a decentralized approach based on information exchange events into a centralized approach based on central data repository. A centralized approach is more convenient for applications that require aggregated information as it eliminates the need to read several events or to execute distributed queries over multiple devices.

Applications and services dealing with context information and devices can be developed at any level of the architecture. Applications can use device services at syntactic or semantic level or receive events with information updates using any abstraction bus. Moreover, applications can formulate queries over the global model in the context server.

3.3 Information Abstraction and Modeling

This section describes how the proposed architecture abstracts information coming from hardware devices like sensors. The following example shows how the state of a magnetic contact attached to a door is converted to the state of the door. The process involves the following steps:

– The *Device Driver* receives a message from the home automation network that indicates that the magnetic contact has the value 1 (i.e. the magnetic contact is open). This message is sent to the *physical abstraction bus* in the same format that is used in the home automation network.
– This event in the *physical abstraction bus* is received by the *device service* associated to the magnetic contact. The *device service* converts the physical message to an XML-based message that represents the state of any magnetic contact regardless of its model or manufacturer. This XML-based message is sent to the *Syntactic abstraction bus* as an event.
– The *device semantic service* associated to the magnetic contact receives the XML-based event from the *Syntactic abstraction bus* and uses the meta-information of the device to convert the event. The meta-information indicates that the magnetic contact is attached to the home's main door and the relationship between the contact state and the door state. Using this meta-information the *device semantic service* generates an RDF fragment according to the system ontology that indicates that the door is open. This RDF fragment is sent to the *semantic abstraction bus* as an event.
– Finally, the event sent to the *Semantic abstraction bus* is received by the *context server*. The *context server* updates the global model with the new information in order to maintain an up-to-date model of the physical world.

4 Prototype

The proposed architecture has been implemented using the OSGi Service Platform (www.osgi.org). For this work Eclipse Equinox, which is a full implementation of the OSGi R4 specification, has been used (www.eclipse.org/equinox). Each entity of the architecture is implemented as an OSGi service and the different communication buses are based on the Event Admin service (a system service included in the OSGi standard).

The generation of the XML-based events for the *syntactic abstraction bus* is based on the XML serialization of Bean Objects included in Java (the java.mbean API). In this way XML messages are generated by serializing the Java objects that implement *Device Services*.

The messages that are sent to the *semantic abstraction bus* are generated by a simple mechanism based on templates. *Device Semantic Services* generate RDF fragments from RDF files with special marks (templates). As there is not yet an existing standard ontology to represent devices and context information [10], the system allows the developer to change the system ontology and the RDF fragments generated by the *device semantic services*. Instead of developing a specific ontology for the architecture, the ontology developed in the AMIGO project [6] is used in the prototype as it is publicly available.

The *context server* service is implemented using the Jena RDF API library (http://jena.sourceforge.net). Although there are other options to work with RDF and ontologies, Jena is the most widely used library and it has a large number of functions. Using Jena, the *context server* implements the central repository of information in RDF format that can be queried using the SPARL query language for RDF.

Currently, the prototype has more than 70 bundles (OSGi name for each individual software component), including some example applications, several device drivers, and more than 20000 code lines.

5 Evaluation

The evaluation and validation methodology has taken into account three different aspects, which are presented in the following sections.

5.1 Device Integration

To verify the feasibility of integrating devices available on the market, the following devices have been successfully integrated in the prototype by implementing specific *device drivers*: Home automation devices (devices following the BUSing protocol which is similar to X-10 - www.ingeniumsl.com/website/en), Ubisense Location System (www.ubisense.net), Multimedia devices based on the UPnP standard and following the Media Render and Media Server profiles (www.upnp.org) and Sun Spot wireless sensor devices (www.sunspotworld.com).

5.2 Example Applications

To verify the viability of developing different types of applications and to study the versatility and reusability of the architecture, diverse types of example applications have been developed:

- Voice Control System. This application uses a VoiceXML server to build a voice interface that allows users to control home automation devices by using voice commands. The application works at syntactic level because it only needs to deal with the available devices. Although this example application has only about 3,700 code lines, the global number of code lines used is approximately 12,000 as it depends on 18 bundles of the architecture (the bundles of the semantic and physical levels).
- Multimedia follow-me application. This application uses the location information provided by the Ubisense System to show each user a specific photo gallery or music play list in the nearest UPnP device as the user changes his location. Semantic events and the information provided by the context server are used to calculate the position of each device and to acquire users' positions. The information used is represented in RDF because the application works at the semantic level. This application has only 2,000 code lines but it uses 17,000 as it depends on the bundles of the physical, syntactic and semantic levels.
- Contextual security application. This application defines a set of rules that applies security restrictions proactively by taken context conditions into account. For example, all the power points are disabled when a child is alone in a room. The application uses a rule reasoning service implemented over the *context server* global model based on the general purpose rule engine included in Jena. Thus the application has been implemented as a set of rules following the syntax of the Jena rules and the ontology used at semantic level (i.e. OSGi or Java knowledge is not necessary). This application has only 138 code lines but it uses 18,000 as it depends on all the bundles of the system.

All these applications have been developed and tested with real hardware in the ambient intelligent laboratory of Fundación CTIC. Results show that applications in higher levels need less code lines and are more hardware independent. Moreover the use of well known technologies such as Java, OSGi, SOA or RDF enables a faster learning curve for developers.

5.3 Performance Analysis

Performance experiments were conducted to study the response time of the system under different conditions and to analyze the viability of deploying the prototype in a real environment. The performance experiments were developed using simulated models, simulated devices and the developed prototype. The number of devices in the system was increased from 1 to 1000 in the experiments (1000 devices is the number of devices of a medium-size hotel, much higher than the number of devices in a common home - about 20). After instantiating a specific number of devices in the system, a software packet generator was used to simulate bus messages coming from hardware devices. The experiments were conducted in a PC with an Intel Core Duo at 1.66 GHz processor, 2 GB of memory, Windows XP operating system and Java 1.6.0_12.

Figure 2 show the memory consumption and the time needed to instantiate the devices when the context server of the system is enabled and disabled. The memory consumption is less than 50 MB when the number of devices is maximum and the context server is enabled. This memory consumption is acceptable even for an embedded device although it is clear that the global model of the context server could cause an out-of-memory problem if the number of devices were too high. Although the instantiation time is high, this time is only a penalization when the system starts.

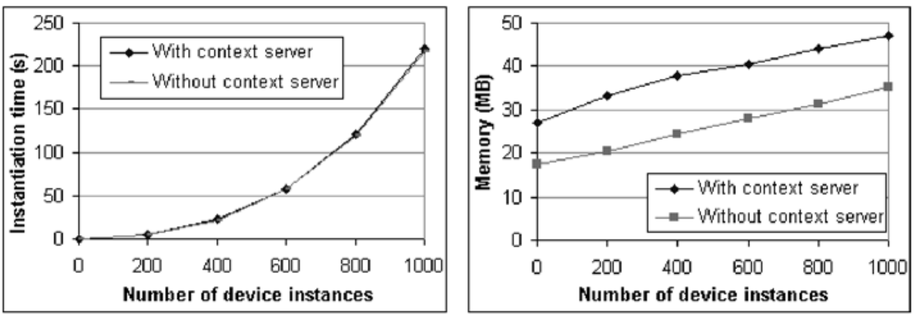


Fig. 2. Instantiation time (*left*) and memory consumption (*right*)

Figure 3 show the time that a message coming from a hardware device needs in order to go from the physical abstraction bus to the semantic abstraction bus or the context server. This time represents the time that the systems needs to abstract context information. These results were obtained when the message rate is about 50 messages per second. This rate is the maximum packet rate in a typical home automation network such as KNX.

The response time of the architecture is acceptable (less than 5 milliseconds in all cases), and stable (the response time does not change as the number of devices increases). Results show that the context server is the slowest element in the system due to the fact that maintaining a non-reasoned RDF model in memory is expensive (with 1000 devices the model has 16326 triples).

Although the architecture allows the implementation of applications using a centralized approach or an event-driven approach, the performance of the context server could impose a limitation in the centralized approach. Depending on the size of the centralized model in the context server, applications with hard temporal constraints could not be implemented properly.

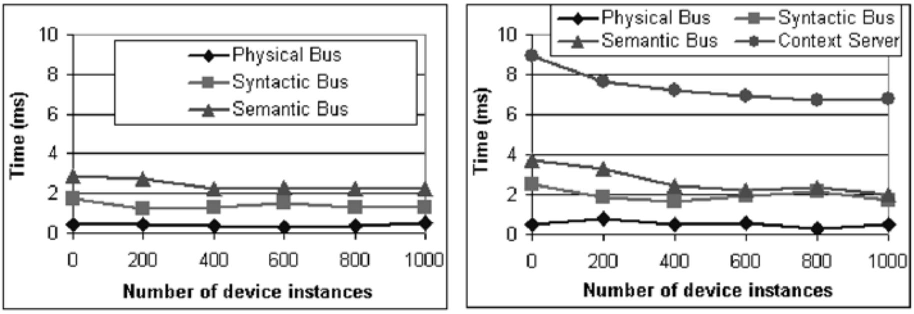


Fig. 3. Abstraction time: without context server (*left*) and with context server (*right*)

6 Related Works

DOG [8] proposed an architecture that represents devices as ontology instances and the interaction with the devices is carried out using messages at the ontology level. The overhead of the central repository with the ontology instances and the use of semantic technologies is not addressed. This paper proposes a hybrid system that represents devices not only as ontology instances, but also as services. This dual representation allows the use of the architecture without a central element (i.e. the context server can be disabled).

Lopez de Ipiña [9] presents an architecture that represents devices as services and uses a rule-based inference system to abstract context information. The rule-based inference system runs in a central element, which has a global model of the system represented using an ontology. A performance evaluation of this system shows unacceptable times for the execution of the rule-based inferences in a real environment. The performance results presented in this paper represent an improvement over de Ipiña's system because in this architecture each individual service representing a hardware device carries out the context information abstraction using a template mechanism.

Gu [7] proposes an architecture based on a central element which executes ontology-based reasoning tasks to abstract context information using a global model. As in the work of Lopez de Ipiña, the performance of the inference process is unacceptable when the number of represented entities is high. Moreover, the number of represented entities in the works of Lopez de Ipiña (up to 100 devices) and Gu (up to 100 instances) is lower than the number of entities used in the performance experiments presented in this work (up to 1000 devices).

Finally, this is the only architecture in which the user can disable the semantic layer or replace easily the ontology due to the template system used in semantic services. The user can choose between the higher abstraction and overhead of the semantic layer or the lower abstraction but faster runtime of the syntactic level.

7 Conclusions and Future Work

This work presented a novel architecture enabling applications to manage devices in smart homes and to use abstract context information acquired with them. The architecture is based on the SOA paradigm and a multi-level structure that allows the integration of different types of devices and the development of different types of applications. It simplifies the development of applications that follow the principles of the ambient intelligent paradigm and that are independent of the underlying physical infrastructure. The evaluation carried out shows acceptable performance that will allow the deployment of the architecture in a real environment overcoming the main problems of other related works. The modular design of the architecture allows the implementation of applications with different temporal constrains and different levels of abstraction.

Future work will include the integration of more types of devices, deployment in a embedded device, and the distribution of the architecture and the context server. Long-term work will include the liberation of the architecture and its implementation under an Open-Source license in order to allow the scientific community to compare, reproduce and verify the presented results.

References

1. Cook, D.J., Das, S.K.: How smart are our environments? An updated look at the state of the art. Pervasive and Mobile Computing 3(2), 53–73 (2007)
2. Augusto, J.C., Nakashima, H., Aghajan, H.: Ambient Intelligence and Smart Environments: A State of the Art. In: Handbook of Ambient Intelligence and Smart Environments, pp. 3–31 (2009)
3. Chan, M., Estéve, D., Escriba, C., Campo, E.: A review of smart homes - Present state and future challenges. Computer Methods and Programs in Biomedicine (2008)
4. Chen, C., Helal, S.: Sifting Through the Jungle of Sensor Standards. IEEE Pervasive Computing 7(4), 84–88 (2008)
5. Edwards, W.K., Grinter, R.E.: At home with ubiquitous computing: Seven challenges. LNCS, pp. 256–272 (2001)
6. Magerkurth, C., Etter, R., et al.: An intelligent user service architecture for networked home environments. In: 2nd IET International Conference on Intelligent Environments, IE 2006 (2006)
7. Gu, T., Pung, H.K., Zhang, D.Q.: A service-oriented middleware for building context-aware services. Journal of Network and Computer Applications 28(1), 1–18 (2005)
8. Bonino, D., Castellina, E., Corno, F.: The DOG gateway: enabling ontology-based intelligent domotic environments. IEEE Transactions on Consumer Electronics 54(4), 1656–1664 (2008)
9. López-de-Ipiña, D., Almeida, A., et al.: Dynamic discovery and semantic reasoning for next generation intelligent environments. In: IET International Conference on Intelligent Environments (2008)
10. Strang, T., Linnhoff-Popien, C.: A context modeling survey. In: Workshop on Advanced Context Modelling, Reasoning and Management as part of UbiComp (2004)

A Distributed Many-Camera System for Multi-person Tracking

Claus Lenz[1], Thorsten Röder[1], Martin Eggers[2], Sikandar Amin[2], Thomas Kisler[2], Bernd Radig[2], Giorgio Panin[1], and Alois Knoll[1]

[1] Robotics and Embedded Systems Lab, TUM
{lenz,roeder,panin,knoll}@in.tum.de
[2] Intelligent Autonomous Systems Group, TUM
{eggers,amin,kisler,radig}@in.tum.de

Abstract. This article presents a modular, distributed and scalable many-camera system designed towards tracking multiple people simultaneously in a natural human-robot interaction scenario set in an apartment mock-up. The described system employs 40 high-resolution cameras networked to 15 computers, redundantly covering an area of approximately 100 square meters. The unique scale and set-up of the system require novel approaches for vision-based tracking, especially with respect to the transfer of targets between the different tracking processes while preserving the target identities. We propose an integrated approach to cope with these challenges, and focus on the system architecture, the target information management, the calibration of the cameras and the applied tracking methodologies themselves.

1 Introduction

1.1 Related Work

Intelligent camera surveillance is commonly employed for both security purposes as well as for smart rooms, which can autonomously react to perceived situations. Such systems can be found operating in real-time or focusing on the post-processing of previously acquired video data. Several surveys including Valera et al. [13] (with an emphasis on distributed systems) and Šegvić et al. [11] give a good overview of state-of-the-art techniques in that field.

A multi-agent-based approach is presented by Patricio et al. [7]. Smart rooms also frequently employ visual tracking, such as Lanz et al. [4]. Teixera et al. [12] present a camera sensor network for behavior recognition using address-event image sensors and sensory grammars in an assisted living environment. Other powerful approaches using smart-cameras with onboard processing that directly deliver data instead of images are presented by Rinner and Wolf [9] or in Hengstler et al. [3] with an eye on application oriented design of the sensor network. A related approach, also using color information and Monte-Carlo filtering using distributed cameras is described by Yamasaki et al. [14].

Precedents prove to be hard to find with respect to the large-scale and the real-time operation as presented in this article.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 217–226, 2010.

1.2 Overview of the System

Apartment Mock-Up. On an experimental area of approximately 100 square meters, a kitchen and a living room have been arranged and furnished accordingly (see Figure 1), in which robots and humans can move freely. If we consider the exemplary scenario of robots serving beverages to humans, several challenges arise due to the fact that different people can order specific drinks. Therefore, a robust and fast way to estimate the position of humans is required, including a unique identification in order to be able to serve the desired beverage to the correct person even if the person moves along while fetching the drink itself. This information can also be useful to facilitate other human-robot joint activities including teaming, task-distribution among robots, or natural motion planning. For a comprehensive report on related research conducted on this mock-up, please refer to Brščić et al. [1].

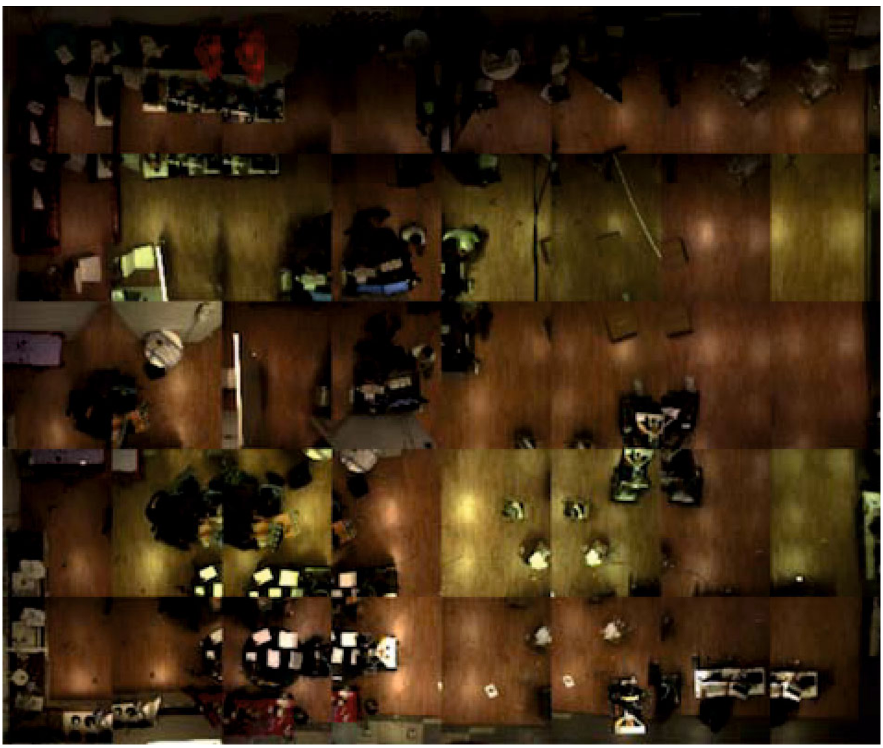


Fig. 1. View of the 40 partly overlapping camera views observing the apartment mock-up. Different illumination conditions are caused by object shadows, while varying chromatic appearances result from using cameras with different specifications and specular reflections are caused by light sources behind the ceiling-mounted cameras.

Hardware Installation. The installed hardware consists of a total of 40 color cameras, each having a native resolution of 1024 × 768 pixels at a rate of 28 to 30 frames per second. All cameras are Ethernet-connected using the GigE-Vision communication standard as described in [2], and are installed on a metal scaffolding mounted at the ceiling. The camera fields of view (FOVs) cover the whole area facing top-down. This setup achieves a total FOV redundancy of approximately 75% at a height of 1.7 m, which according to Ogden et al. [5] is the approximate average height of an adult person. The cameras are grouped in threes and pairs respectively to form 14 camera groups, each of which is in turn linked via a Gigabit Ethernet (GigE) switch to a diskless processing node, where image capturing and processing itself takes place. A single server computer manages the diskless node network, and hosts the server applications described in Section 3.2.

For robustness and load-balancing reasons, adjacent cameras are assigned to different camera groups. This helps to compensate for the observable fact that human beings tend to flock together in social scenarios, rather than distribute evenly over the surveyed area. Besides the drawbacks of a relatively high amount of cameras being required to cover the area, and the requirement of managing frequent transfers of targets between camera FOVs, this specific camera setup offers the following advantages:

- The image resolution of each local FOV shows a high quality, compared to solutions using less cameras.
- All cameras are almost co-planar. This allows for application of the same tracking techniques and assumptions for the whole covered area, which reduces the system and algorithmic complexity.
- Under the assumption that people do not climb over each other, a full-body mutual occlusion is almost impossible to occur.
- Since the camera transformations w.r.t. a common world frame are known, the hand-over regions (shared FOV regions in which a target transfer may occur) can be defined and evaluated easily.
- Because of the local distribution of computational power, the number of simultaneously tracked people can be increased or the computational power can be shared in a scalable and distributed way with other computationally intensive approaches e.g. for gesture recognition or or activity analysis.

2 Single Camera Person Tracker

2.1 Person Detection

New persons are detected as they enter the FOV of a camera using a foreground-background segmentation (adaptive mixture of Gaussians) [15], followed by a blob clustering [10] and a data association method. No further a priori information about the person's appearance including color or texture clues is used up to now. Every blob that results from a foreground region is classified as a human by analyzing the area size and the aspect ratio of the outer dimensions.

The center of mass of a blob as x and y position on the sensor projection layer is subsequently upgraded to a 3D translatorial pose of the human target in the world frame using the extrinsic camera parameters. A virtual ray given by the focal point of the camera and the computed mass center of the blob is casted and then intersected with the ground floor, which has a z-coordinate of 0. Minor position errors due to perspective distortions and the fact that the height of the detected person can not be estimated using top-down mounted cameras are corrected in the first tracking step.

In addition, the person detection method is employed for validation and target association during the target transfer process described in Section 3.2.

2.2 Model Building

The 3D model approximating the human shape consists of a simple cylinder, roughly corresponding to human size in real-world coordinates. Once the overall scale has been computed from the detected blob, the relative proportions (location of the head with respect to the torso, relative size, etc.) are fixed according to an average model of the human body. The statistical color model is obtained by collecting the image pixels for the respective blob, in a 3D histogram in HSV color space. Different bin sizes are used, in order to give more importance to the color attributes rather than the intensity values, which are illumination-dependent. In this case, a robust combination of 8 bins for hue, 8 bins for saturation and 4 bins for value were used. In order to collect only pixels that do belong to the person, the background segmentation image is treated as a mask.

2.3 Multi-target MCMC Filter

The color model of a person is now used in order to instantiate a new *target*, to which a unique ID number is assigned, and that will be tracked across the image sequence by this camera, until the person leaves the camera's FOV.

Tracking operates on a pre-defined set of degrees of freedom, which for our rigid 3D shape model is defined solely by the 2D translation on the floor (x, y coordinates). Therefore, the state vector of the $i - th$ target is given by $s^i = (t^i_x, t^i_y)$.

The tracking methodology basically consists in matching the reference color histogram to the current image, underlying the projected shape of the person. By using a calibrated camera model, we also take into account perspective effects while computing the *silhouette* in the camera image. These effects have a high impact for our setup, since the relative distance between the camera and the person is comparable with the depth extension of the target (i.e. the height of the person). Therefore, they cannot be neglected, especially for people in the peripheral view field.

In order to estimate the state of each person on the image sequence, a Bayesian Monte-Carlo tracking approach is used, described in more detail in [6]. This methodology consists of a particle filter, which maintains the global system state,

$s = (s^1, ..., s^m)$ of the m currently active targets in the scene, by means of a set of hypotheses s_h (or *particles*), that are updated from frame to frame by means of Markov-Chain Monte-Carlo sampling (MCMC). In particular, the Markov-Chain generation proceeds by iterating (for each particle $n = 1, ..., N$) two steps, that correspond to the *Metropolis-Hastings* algorithm.

The efficiency of the MCMC formulation is due to the fact that we update a *single target* i at a time (randomly chosen), which results in the computation of the proposal ratio only for this target $P\left(s_{i,t} \middle| s_{i,t-1}\right)$.

3 Upgrading to a Many-Camera Set-Up

3.1 Calibration of Cameras

The calibration procedure is performed in two steps, with the intrinsic camera parameters being determined independently on all cameras in the first step. The second step then aims at determining the extrinsic parameters of all cameras w.r.t. the global world origin. To ensure optimum inter-camera consistency of calibration, which is an important issue regarding the transfer of tracked persons from one camera to another, we make additional use of an infrared tracking system consisting of six Visualeyez VZ 4000 [8] tracker bars, that measure accurately with a root mean square error of below 0.5 mm. A standard calibration pattern was enhanced with 4 infrared diodes, so that the infrared tracking system is able to determine the pose of the calibration pattern accurately at any position within the scene. The obtained pose information was transmitted to the diskless clients managing the cameras, with the cameras being used to record image sequences simultaneously, which can then be used to determine the extrinsic parameters via prevalent calibration methods.

3.2 Management of Tracking

The tracking algorithm described in Section 2 is working independently and asynchronously for each camera without knowledge about other cameras or synchronization mechanisms on the client sides. Using this design principle, a server application handles the centralized management of the tracking results and takes care of the transfer of tracked targets between camera FOVs (and consequently, processing nodes). The advantages of this approach lie with its full scalability and the lack of need for synchronization. To realize the approach, global modules for registration of the single trackers, a global display module, a module for the generation of unique target IDs and for the management of the transfer of targets between tracking clients were implemented. All of these modules are running on the server computer of our hardware installation, while the client applications, responsible for detection and tracking, run on the diskless processing nodes, as described in Section 1.2.

Global Registration. The global registration instance of the system possesses a global representation of all connected cameras and processing nodes. This includes knowledge about 1) the unique IDs of the cameras and their connection to specific processing nodes, 2) the extrinsic and intrinsic parameters of each camera, 3) the possibility to enter the surveyed area through a specific camera FOV, 4) the expected sizes for camera images for the streaming process and 5) the expected size of a connected display.

The aforementioned knowledge about the setup is loaded to the system and can be exchanged, in order to adapt the system to other scenarios and setups. The client applications, which register their camera in the registration server, obtain the expected size of images to be sent to connected displays in return. Furthermore, the client applications request the global knowledge about the surveyed area and use the obtained information to instantiate their system with the correct intrinsic and extrinsic camera parameters. This leads to the advantage that no client application has to keep local information on the setup. Therefore, if the setup is changed e.g. by adding new cameras, moving cameras or recalibration, a simple restart of the registration and the client applications updates the whole system.

Global ID Generation. In a distributed, asynchronous tracking system with completely independent tracking processes, it is an essential need that the tracked targets keep their identity after a target transfer from one tracking process to another, which occurs if a tracked target switches between camera FOVs. Once the detector module of a client application has found a new target, it can only be introduced and added in the server to the global system state using a uniquely generated ID. Therefore, the client *requests* a new ID from the global ID generator module. Using this ID, the client is able to add the target in the server with its current position in world space. Subsequently, the target's position will be broadcast to connected processes, e.g. those running on the robots in the scenario described above.

All clients have a direct connection to the server to add targets, update the target positions, or to initiate handover targets in case they are leaving the surveyed area. In order to maintain the scalability of the system, the server module does not possess a priori knowledge on the number of client applications. Therefore, a communication channel was created that broadcasts the control commands: a) transfer target: to give a client the responsibility to track a certain target and b) remove target: to release the responsibility of a client to track a certain target. Every client listens to this control channel and reacts only if the mentioned target is within his responsibility (remove target) or if he should take over the tracking (transfer target).

Management Server. A challenging task is to decide which client should do the tracking, when to transfer targets, and where to transfer targets to. Performing the evaluation of 40 camera location at every position update (up to 15 Hz) requires very high computational power ($n-1$ comparisons). At this point, it is possible to utilize the extrinsic and intrinsic calibration parameters of the cameras, that are stored in the registration server. After fetching this information, a quad-tree is built up, dividing the scene in quarters: In the first

(a) Frame 100 (b) Frame 150 (c) Frame 200 (d) Frame 250

Fig. 2. Example image sequence: joint view of four adjacent cameras tracking two different persons. a) and (b) show the tracking of a single person before and after a hand-over of the person target took place. In the same way (c) and (d) depict the tracking results of two persons, before and after the respective handover situations have occurred.

step each center point and FOV of all cameras are projected on the floor. Now, the tree is iteratively refined dividing each node in 4 sub-nodes (areal quarters) until each node has only one camera left.

Fig. 3. Target transfer example sequence: Joint view of four adjacent cameras. Areas marked by blue dots represent valid target transfer regions between adjacent FOVs. The dots result from the application of space discrete (in world coordinates) tests, which consist of evaluation of the target transfer tree that was pre-computed from the respective camera projection matrix. The red cylinder represents the tracked person that walks through adjacent FOVs.

The tree nodes contain all the important information to select the optimal camera to transfer a target to. This includes the projection center of the camera on the ground floor, the FOV, the camera index, and the hostname of the client to which the camera is connected. After each update of the target's position, the pre-calculated tree can be traversed efficiently to find the next host and the next camera index. Figure 3 depicts 4 adjacent cameras with blue dots which represent valid target transfer points between cameras. These points are the result of a space discrete (in world coordinates) test using the pre-calculated target transfer tree.

4 Results

The system was implemented using the proposed tracking and communication architecture. Subsequently, the system was evaluated by tracking persons exhibiting high variance w.r.t. their height, appearance and motion habits, as well as under different illumination conditions during the respective days. This was done in order to test the applicability of the rather coarse model assumptions, being the average height of people (1.7m), the applicability of the rigid non-articulated cylindrical shape representation and the evaluation of color statistics from the top view.

As depicted in Figure 1, the client-server architecture is able to stream miniaturized camera images to multiple connected display processes, which may be distributed over the internet. Figure 2 depicts the joint view of four adjacent cameras simultaneously tracking two individual persons. (a) and (b) shows the tracking of one person including a hand-over procedure. (c) and (d) depict the tracking of two persons. Figure 3 illustrates the pre-computed target transfer regions in a joint view of four adjacent cameras. The transfer regions can be estimated by evaluating the quad-tree which is computed using the extrinsic camera parameters.

Regarding the the coarse 3D cylindrical person model, several tracking approaches were tested. E.g. Figure 4 shows the comparison of two different evaluation strategies for the sampling of the color statistics, comparing the accuracy achieved and required computational time. The first strategy consists of the sampling of pixels within the underlying rotated bounding box of the cylindrical model. The second strategy projects the 3D cylindrical model into the sensor image and thus computes the full and exact shadow, which is used as a mask before pixel sampling takes place. Based on these results, we decided to adopt the full shadow approach, since the improved accuracy outweighs the drawback of a slightly higher computational cost in our setup.

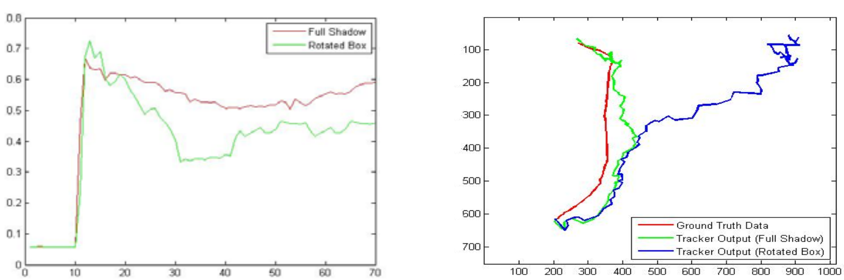


Fig. 4. Using the full shadow of the cylindric human shape approximation to compute the measurement results in a significant gain in accuracy while requiring only slightly more computation time. *left*: Time needed to evaluate measurement using rotated box or the full shadow. X axis denotes the number of frames, Y axis denotes the time in seconds, running three tracking modules simultaneously. *right*: Accuracy using rotated box or full shadow. X and Y axes denote pixel positions within the image.

5 Conclusion and Future Work

In this paper, a flexible, scalable and modular many-camera system for simultaneous tracking of multiple persons using natural features was presented. The approach was realized and evaluated using an apartment mock-up, sensorized by 40 GigE cameras which fully cover its approximately 100 square meters. Given this large amount of cameras, distribution of the computational processing among multiple computers is required, which is addressed using 14 diskless client processing nodes operating up to three cameras each. A functional system for the management of target detection, target tracking and target transfer between processing nodes was presented.

Future work includes improving the detection step by application of a more robust classification of persons, rather than assuming every foreground pixel to be belonging to a human. This will help avoid erroneous detection alarms. Furthermore, the degrees of freedom tracked can be extended by including the estimation of persons' rotation angles. In our application context, this additional information would allow robots to intentionally approach persons from specific directions, as well as facilitating the evaluation of psychological experiments within the surveyed area.

Acknowledgements

This work was partly supported by the cluster of excellence *CoTeSys – Cognition for Technical System* (http://www.cotesys.org) within the projects *ITrackU – Image-based Tracking and Understanding* and *MuJoA – Multi Joint Action* funded by the DFG.

References

1. Brščić, D., Eggers, M., Rohrmüller, F., Kourakos, O., Sosnowski, S., Althoff, D., Lawitzky, M., Mörtl, A., Rambow, M., Koropouli, V., Hernández, J.M., Zang, X., Wang, W., Wollherr, D., Kühnlenz, K., Mayer, C., Kruse, T., Kirsch, A., Blume, J., Bannat, A., Rehrl, T., Wallhoff, F., Lorenz, T., Basili, P., Lenz, C., Röder, T., Panin, G., Maier, W., Hirche, S., Buss, M., Beetz, M., Radig, B., Schubö, A., Glasauer, S., Knoll, A., Steinbach, E.: Multi Joint Action in CoTeSys - setup and challenges. Tech. Rep. CoTeSys-TR-10-01, CoTeSys Cluster of Excelence: Technische Universität München & Ludwig-Maximilians-Universität München, Munich, Germany (June 2010)
2. Chamberlain, G.: GigE Vision: standard route to video over IP. Industrial Ethernet Book (33), 35 (July 2006)
3. Hengstler, S., Aghajan, H., Goldsmith, A.: Application-Oriented Design of Smart Camera Networks. In: First ACM/IEEE International Conference on Distributed Smart Cameras, ICDSC 2007, pp. 12–19 (2007)
4. Lanz, O., Chippendale, P., Brunelli, R.: An appearance-based particle filter for visual tracking in smart rooms (2007)

5. Ogden, C., Fryar, C., Carroll, M., Flegal, K.: Mean body weight, height, and body mass index, united states 1960–2002. Advance data from vital and health statistics (347) (2004), http://www.cdc.gov/nchs/data/ad/ad347.pdf
6. Panin, G., Röder, T., Knoll, A.: Integrating robust likelihoods with monte-carlo filters for multi-target tracking. In: International Workshop on Vision, Modeling and Visualization (VMV), Konstanz, Germany (October 2008)
7. Patricio, M.A., Carbo, J., Perez, O., Garcia, J., Molina, J.: Multi-agent framework in visual sensor networks. EURASIP Journal on Advances in Signal Processing(Print) 2007(7) (2007)
8. Visualeyez VZ 4000 (2009), http://www.ptiphoenix.com/VZmodels.php
9. Rinner, B., Wolf, W.: An introduction to distributed smart cameras. Proceedings of the IEEE 96(10), 1565–1575 (2008)
10. Rocha, L., Velho, L., Carvalho, P.C.P.: Motion reconstruction using moments analysis. In: SIBGRAPI 2004: Proceedings of the Computer Graphics and Image Processing, XVII Brazilian Symposium, pp. 354–361. IEEE Computer Society, Washington (2004)
11. Segvic, S., Ribaric, S.: A software architecture for distributed visual tracking in a global vision localization system, p. 365 (2003)
12. Teixeira, T., Lymberopoulos, D., Culurciello, E., Aloimonos, Y., Savvides, A.: A lightweight camera sensor network operating on symbolic information. In: Proceedings of 1st Workshop on Distributed Smart Cameras, Held in Conjunction with ACM SenSys. Citeseer (2006)
13. Valera, M., Velastin, S.: Intelligent distributed surveillance systems: a review. IEEE Proceedings Vision, Image and Signal Processing 152(2), 192–204 (2005)
14. Yamasaki, T., Nishioka, Y., Aizawa, K.: Interactive retrieval for multi-camera surveillance systems featuring spatio-temporal summarization. In: MM 2008: Proceeding of the 16th ACM International Conference on Multimedia, pp. 797–800. ACM, New York (2008)
15. Zivkovic, Z.: Improved adaptive Gaussian mixture model for background subtraction. In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 2 (2004)

An Open Distributed Framework for Adaptive User Interaction in Ambient Intelligence

Mohammad-Reza Tazari

Fraunhofer-IGD, Darmstadt, Germany
saied.tazari@igd.fraunhofer.de

Abstract. Challenges of handling user interaction in Ambient Intelligence environments are manifold. The systems installed in these environments are highly distributed with dynamic configurations in terms of integrated devices and installed applications. Context-awareness, personalization, and multimodality are critical for supporting more natural interaction and optimizing the interaction in an adaptive way. Research activities have dealt with different specific problems in the field and now it is high time for moving towards an open framework with a more comprehensive solution. This paper introduces results of such work with a high level of freedom in developing and deploying applications without needing to care about the available I/O infrastructure. The other way around, the latter can be changed without worrying about the application side. The independence of applications from the available I/O infrastructure helps to share mechanisms and manage such a complex scene more adequately. The key idea behind this framework is the natural distribution of tasks according to the real scene using a middleware solution supporting seamless connectivity and goal-based interoperability.

1 Introduction

Ambient Intelligence (AmI) [10] is the European initiative to promote the creation of smart and reactive environments that empower humans to easily utilize the services available and assist them in carrying out their tasks and activities in that environment. Long enough explicit user interaction stayed in AmI environments in the shadow of "implicit interaction". To satisfy the core requirement that the system reacts to the contextual changes automatically, researchers were focued on changes in the user context and at most those user actions in the environment that do not address the "system" explicitly [13], [16], [11]. However, with the proliferation of (multi-)touch sensing, big displays as well as small displays embedded in all possible devices, and new interaction forms supported by special devices (e.g. WiiMote) on one side, and progresses in speech recognition, natural language processing, and gesture recognition on the other side, explicit user interaction in AmI environments is becoming more and more important.

This paper presents the results of work on an open distributed framework for adaptive user interaction in AmI environments that has been carried out within the EU IST project PERSONA with a focus on the explicit interaction with elderly people possibly afflicted by some impairments.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 227–238, 2010.

Treating AmI environments as open distributed systems is somewhat natural [8], which would suggest to separate the application layer from the presentation layer. That is, arbitrary applications must be able to utilize the same hardware configuration with several displays, loudspeakers, microphones, . . . distributed in the environment. This means that applications should be able to describe their dialogs in a modality- and layout-independent way and communicate such descriptions to a UI framework using well-defined protocols and interfaces, delegating the selection of the right modality and output devices to the UI framework. As a result, applications could be plugged in/out according to user needs while sharing the available I/O infrastructure. The other way around, the UI framework must be designed in a way that components managing the I/O channels (from now on: *I/O handlers*) can be added, removed or replaced freely, e.g. in accordance with the enhancement, discontinuation or improvement of the available I/O channels, or simply because a certain alternative handler is preferred. Again, well-defined protocols and interfaces are needed to adequately support the free plug-in / -out of I/O handlers. Figure 1 summarizes the scene based on the two first fundamental requirements by abstracting the UI framework as a black box. Another fundamental requirement is the support for adaptive interaction.

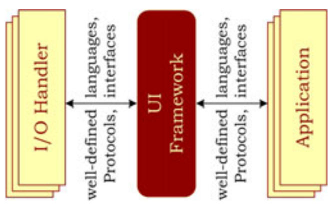


Fig. 1. A UI framework separating the application & presentation layers

The selection of appropriate modalities and devices often depends on the user context, capabilities, and preferences. After that, and at the time of "rendering" and determining the "layout" by the I/O handlers, context-, personalization- and dialog-related parameters can similarly be used to optimize the outcome from the view point of the addressed user, e.g., by incorporating the user preferences regarding the voice gender and form of address in speech-based interaction as well as the adaptation of volume level to the situation. Even applications may consider such parameters in the process of preparing the content to be presented to the user. E.g., adaptive applications may choose the content language or prioritize (e.g., through context-aware sorting) or emphasize different portions of info differently. Hence, the following adaptation-related requirements will result:

- The UI framework must provide applications with facilities to easily access relevant context and personalization parameters that may affect the process of preparing the content.
- Along with dialog descriptions, applications must provide the UI framework with content-related info relevant for the presentation, such as the language and privacy level of the content.

- The UI framework itself must select the appropriate modalities and devices for each dialog in a context-aware and personalized way.
- The UI framework must instruct a selected I/O handler how to present a dialog by passing an appropriate set of adaptation parameters to it – e.g., voice gender and volume level from the above example.
- Upon changes in the user context, the UI framework must redo the previous two steps for the running dialogs and either change the instructions to the previously selected I/O handler or switch to an alternative I/O handler (cf. UI migration in [1]); as a result, the UI framework will be able to seamlessly support dynamic adaptation and, for example, implement "follow-me" scenarios (i.e., the user changes the location and immediately the I/O channels available at the new location are used for continuing the dialog that was not finished yet) or react in a privacy-aware way (e.g., transfer the dialog to more private channels when someone enters the same location as the interacting user and there is a conflict with the privacy-level of the presented content).

Last but not least, in a distributed environment with several I/O channels, it is important to pay attention to the fusion and fission of modalities on the input and output side, respectively:

- With the above understanding of I/O handlers, they can be seen as entities responsible for interpreting user input coming from input channels monitored by them. The interpretation process may need to combine input coming from different channels based on some synchronization mechanism; e.g., when a user says "turn off the lamp over there" while pointing with a finger to the relevant lamp, only an I/O handler that listens to events provided by both microphones and gesture recognition components can locate the lamp and provide a meaningful command to the system.
- Explicit user input may be provided to an AmI system in the context of a previous system output or in a context-free way. The determination of the type of input (context-free or not) can be another task of I/O handlers. An I/O handler already presenting some system output and waiting for related input can easily match the user input with the expected input if it is provided through input channels monitored by that I/O handler. Otherwise such an I/O handler may remain in the wait state forever while another I/O handler infers a context-free input. Thus, it is important that there are mechanisms to avoid such situations.
- If the output generated by an application and published for presentation to a human user is layout- and modality-neutral, then it is the task of I/O handlers to "render" and present it to the user by utilizing appropriate functions of relevant output devices. "Relevant" in this context has to do with the capabilities of the I/O handler regarding support of different modalities and / or several devices of the same type. The rendering process by a multimodal I/O handler, however, must be based on decomposition strategies – in order to decide which parts of an abstract dialog should be rendered using which modalities – and use synchronization mechanisms between supported modalities. The utilization of output devices may be based on low-level protocols,

mid-level abstractions, such as those provided by a windowing system, or high-level abstractions, such as services. In some cases, it could be necessary to install the same I/O handler on several nodes[1] and distinguish between them using properties, such as installation location.
– To allow multimodal output in absence of multimodal I/O handlers, the UI framework may also provide its own decomposition strategies and synchronization mechanisms.

2 Related Work

A survey of the state-of-the-art shows that many researchers are working on different aspects of user interaction in AmI environments, mostly with a focus on adaptivity and multimodality. In [9], for instance, Flippo et al. present a framework for rapid development of multimodal interfaces (or multimodal I/O handlers). While supporting modality fusion on the input side, the solution disregards modality fission on the output side. As a matter of fact, it needs the application logic to be fed into the framework and, furthermore, does not have distributed environments in focus. A similar solution in its essence has been introduced in [2]; an interesting result of the analysis done in this paper is that access to I/O channels of devices is not posing any requirement beyond the general interoperability requirements of *eventing* and *service call*.

Another track of research is focused on frameworks for content adaptation. The utilization of a distributed infrastructure of I/O channels is normally out of the scope of such solutions. The framework introduced in [7] is an example hereof. A disadvantage of this approach concerns the codifying of the environment and user models into the framework that hardens the enhancement of such models.

Calvary et al. have been working on the so-called “plastic user interfaces” (adaptive UI) in distributed environments for several years (cf. [4,5,1,6]). In [4], they introduce a “situation → reaction” model, where adaptive reactions could be one of *change of the device*, *change of the renderer on the same device*, or *instructing the renderer to adapt itself*. The situation recognition was based on transitions between different contexts. This model partially satisfies the above requirements but it seems that in the realization, the separation between the application and presentation layers as well as support for plug-ability were disregarded. This model was enhanced in [5] by allowing more transition possibilities between different contexts. In [1], mechanisms are introduced for managing running dialogs in their original form (by re-calculating the abstract UI from its current concrete form) in order to overcome the problem of migrating concrete UIs. However, migration rules seem to be application-specific instead of being based on an open repository of facts and preferences (the latest work in [6] was more theoretical on identifying the determinants of adaptivity). In all, the solution is based on a centralized situation recognition mechanism that does not

[1] For instance, when a GUI-providing I/O handler uses the windowing system provided by runtime environments, such as Java and Microsoft .Net, in order to avoid complex low-level protocols for monitoring / utilizing the I/O channels.

go for an open UI framework, does not distribute the adaptation tasks among different layers (e.g. application, framework, I/O handlers), and consequently is laid out in terms of devices.

To summarize, our investigations show that there is mostly a focus on specific aspects without needed concepts for a generalized framework respecting the open and distributed nature of AmI systems.

3 The PERSONA UI Framework

The EU-IST project PERSONA aims at providing an open and scalable technological platform that facilitates the development and deployment of a broad range of services in the field of Ambient-Assisted Living (AAL), in PERSONA with a focus on the "smart environment" aspects. AAL is the concept that groups the set of technological solutions, named AAL Services, targeting the extension of the time that elderly and disabled people live independently in their preferred environment. AAL Services provide personalized continuity of care and assistance, dynamically adapted to the individual needs and circumstances of the users throughout their lives. The PERSONA scenarios address the need for social integration and belonging, independence in daily life activities, security and safety at home as well as outdoors, and mobility.

Smart environments are treated in PERSONA as open distributed systems. The open distributed platform of PERSONA should make the formation of smart environments affordable for all. It should be possible to start with a small core and let the system evolve over time so that people can arrange the system according to their individual needs as they arise. A compact piece of software should enable the producers of diverse networking-enabled nodes, such as sensors and controllable gadgets and appliances, as well as developers of applications with different levels of complexity, to produce "PERSONA-aware" components that can be plugged into the system dynamically without much effort. Instead of dealing with several complicated interfaces of different devices, users should experience an integrated world easy to interact with based on natural communication. Finally, the open and distributed nature of the system should lead to more competition on the market promoting the production of not only new components but also improved alternatives for different parts of the existing system, both at the level of application and the platform itself.

3.1 Principles of the PERSONA Middleware

PERSONA abstracts the physical architecture of AAL Spaces (e.g., smart homes of elderly people) as a dynamic ensemble of seamlessly networked nodes. Hence, as a first step the project started to design and develop the smallest piece of software, called the PERSONA middleware, that was needed to make a node "PERSONA-aware" (cf. figure 2). To develop the middleware, a message brokering approach based on ontological match-making has been chosen. The middleware provides four virtual communication buses dedicated to dispatching captured user input

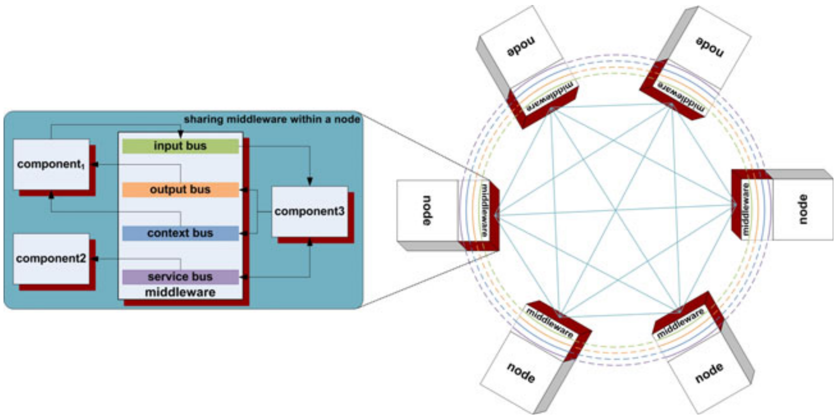


Fig. 2. The PERSONA system as a dynamic ensemble of seamlessly networked nodes

(the input bus) as well as generated system output (the output bus) on one side, and enabling both event-based and call-based inter-component communication (through context and service buses respectively), on the other side. For more info on details of this design principle, please refer to [15].

The PERSONA middleware realizes a reasoning mechanism based on the fundamental concepts of the Semantic Web, e.g. RDF resources and OWL classes, individuals, and class expressions (see www.w3.org/2001/sw/), as well as OWL-S [12]. Two pluggable components that share the same ontology can communicate using the middleware buses while the middleware itself remains neutral in regard to the used ontologies. Data exchanged on the middleware buses is serialized in RDF/XML as soon as it leaves an instance of the middleware and again de-serialized by the receiving instance of the middleware transparently so that members of the buses see the data always in form of objects without needing to deal with XML (de-)serialization.

The buses of the PERSONA middleware introduce certain RDF resources that can be used for exchanging messages on them, and work based on distributed message brokering strategies that provide virtually global views on the buses and hide the distribution of the components on the nodes in the ensemble. The RDF resource used for message exchange on the context bus is called a *context event* [8]. On the service bus basically resources representing a *service profile*, a *service request* or a *service response* can be exchanged [14].

3.2 The Input and Output Buses

The event-based I/O buses of the PERSONA middleware in figure 3 are designed to satisfy the interoperability needs in smart environments related to the explicit interaction between human users and the deployed system. Components with the ability to capture user input publish the related info as input event to the input bus that must find the appropriate applications able to process it. On the other hand, when applications intend to present some info to a user, they publish an

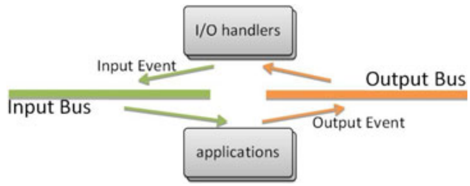


Fig. 3. Communication between applications and I/O handlers

output event to the output bus, relying that the bus will find an appropriate subscriber able to handle the event. Hence, the PERSONA I/O buses are the core part of the UI framework of figure 1 that enable the communication between applications and I/O handlers while avoiding tight connections.

As an application-independent, technological solution, each I/O handler manages a set of I/O channels and subscribes to the output bus by specifying its capabilities, which are used by the output bus in the course of match-making with adaptation parameters. That is, using the adaptation parameters (see section 3.4), the output bus tries to find a best-match I/O handler that receives the content to be presented to the user along with instructions in regard to modality and layout. The selected I/O handler must then convert the application output to the format appropriate for its output channels in accordance with the instructions and monitor its input channels to catch the related user input. Upon recognized input, it must convert it to the appropriate format for the previously handled application output and publish it as an event to the input bus.

When an application that publishes an output event to the output bus expects that related user input is captured and sent back to it, it has to "simultaneously" subscribe to the input bus with the related dialog ID. The same dialog ID must be used in the related input event to be published by the selected I/O handler. The task of the input bus is then as simple as forwarding the received input event to the related subscriber, which is then unsubscribed automatically. For processing context-free user input (cf. section 1), the one special component that subscribes to the input bus at start-up without specifying any dialog ID will always be selected. This special component is called the *Dialog Manager*, which is one of the mandatory components of the PERSONA platform in charge of representing the whole system; e.g., it not only handles context-free input but also provides the system main menu for utilizing available user services. It additionally assists the output bus by providing the adaptation parameters (see section 3.4) and managing a priority queue of published dialogs in order to guarantee that at each point in time only one dialog is presented to each user.

3.3 Dialog Description and User Input Data

The most important part of an output event is the description of the dialog that an application is going to present to a certain user. Such a description must be device-, modality- and layout-independent in order to guarantee a clear separation between application logic and presentation layers of the system. PERSONA relies on XForms specification [3] for this purpose. XForms scripts basically are formed from two parts: an XForms Model specifying what the form is about (the kind of data being collected), and a set of UI controls for presenting the form and collecting input. Investigations showed that the XForms Model is very well compatible with the RDF model. The type system embedded in XForms Model is inherently supported in a more powerful way in PERSONA because of using OWL-based ontologies. This lead to the decision in PERSONA to provide its own dialog package consisting of specific RDF resources representing the equivalents of XForms controls; the data section of XForms is then replaced by RDF resources or OWL individuals originally used by PERSONA applications. Both the form controls and the form data are embedded in a single resource of type "Form" that represents a dialog in the PERSONA system. When an I/O handler finishes a dialog, the data section of the form is automatically added to the input event to be published to the input bus. Two interesting features of the provided dialog package are:

- The data part of the forms is hidden to the I/O handlers and they do not need to do anything for type checking of user input but they simply try to store it using the package facilities. If this attempt fails, they must give some hint to the user using the alert message set by the application with the same semantic as defined by XForms.
- Form controls can be created by the dialog package automatically based on ontological knowledge associated with the data, if desired.

3.4 Adaptivity in the UI Framework

Adaptation in presentation layers has two important determinants: personalization and context-awareness. The PERSONA framework for supporting context-awareness, mainly consisting of the context bus, the Situation Reasoner, and the Context History Entrepôt [8], already covers one of these aspects. A profiling component completes the scene by identifying the personalization parameters and providing mechanisms for dealing with them. In the following, we show how these platform components along with the Dialog Manager mentioned in section 3.2 act as part of the PERSONA UI framework and support the output bus to carry out its share of adaptation tasks (cf. figure 4).

As formerly mentioned, I/O handlers register their profiles to the output bus. The profile consists of items such as appropriateness for certain access impairments, supported languages and modalities, locations where output can be presented, and modality-specific tuning capabilities. On the other hand, applications provide their output using the dialog package sketched in section 3.3, along with content-specific adaptation parameters (APs), such as the addressed user and the content language and privacy level. Any content-specific adaptation, such as a situation-aware sorting, must have been applied to the content previously.

Upon receipt of an output event, the output bus asks the Dialog Manager to add the application-independent set of APs based on the addressed user and the content privacy level provided by the application. These are: the presentation location and modality, access impairments to be considered, and modality-specific

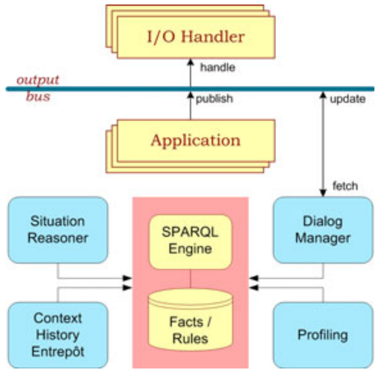


Fig. 4. Components involved in the realization of adaptation mechanism

recommendations. The whole set of APs is then used to choose the most appropriate I/O handler based on an ontological match-making against the registered profiles, and forwarded to the selected I/O handler as adaptation instructions. The I/O handler selected for the presentation of the output must then "render" the modality- and layout-neutral representation of the content, in accordance to the adaptation instructions, in the specified modality under the consideration of the specified language, privacy, and the modality-specific tuning parameters. The result of the conversion must then be presented to the user using an appropriate device at the specified location.

During a dialog is running, the Dialog Manager can notify the output bus when any of the APs changes; the output bus may then either notify the I/O handler in charge of that dialog to redo the above step (if the changes in the APs still match its profile) or switch to another I/O handler (if the new situation cannot be handled by the previous I/O handler). In the latter case, the previous I/O handler is notified to abort the dialog while returning any intermediate user input collected so far, and then the new I/O handler is mandated to continue with the dialog presentation. As a result, follow-me scenarios can be realized as well as switch from a big display to a personal device when other people enter the same location where presentation takes place.

In order to reduce the delay for fetching the APs, the profiling component delegates rules for keeping the APs up-to-date to the Situation Reasoner so that re-calculation and update of these parameters is always running in the background and the Dialog Manager can be sure that the values found in the database are always up-to-date and valid. This way, the fetch process will be as simple as querying the database. As the profiling component and the Context History Entrepôt share the same database for storing their data and the same database is used by the Situation Reasoner for evaluating the rules, powerful rules can be defined that combine both the contextual and the personalization data. These rules are defined in terms of SPARQL (the RDF query language at www.w3.org/TR/rdf-sparql-query) and indexed according to their dependence on database events that should trigger the evaluation of the rule (see also [8]).

4 Conclusions and Future Work

An open distributed framework for adaptive user interaction in Ambient Intelligence was introduced. It is open because of providing well-defined interfaces and protocols that are already disclosed in part. Until the end of the PERSONA project in October 2010 all documentations of these interfaces and protocols will be available on the project Web site in their final versions and the source code of the middleware and many of the platform components is being prepared for publication under Apache License 2.0. It is also open in the sense that it leaves the configuration of AmI systems open allowing for plugging in and out all possible sorts of components, no matter if hardware or software, as long as they follow the specifications. Due to the underlying loose coupling mechanisms, even platform components can be replaced by alternative solutions that fulfill the requirements.

The described framework is distributed not only as a solution for a highly distributed system but also because of its components: *(1)* the middleware is developed as a distributed software system that through cooperation of its different instances can provide a consistent view on the system hiding the distribution of its components; and *(2)*, platform components can be deployed on arbitrary physical nodes in the system due to their full reliance on the middleware for inter-component communication.

Finally, the PERSONA UI framework provides adequate support for adaptivity while letting the user and context models open and pluggable. It does not try to cover all required adaptation mechanisms in a centralized way, but it delegates the tasks to pluggable applications and I/O handlers based on a natural division of work and overtakes only the brokering task that in fact leads to a context-aware and personalized selection of modality and eventually device.

Nevertheless, there are some issues still open. One of the severe issues is related to modality fusion and fission tasks and the related synchronization mechanisms needed: as stated in section 1, although it is reasonable to delegate modality fusion and fission to multimodal I/O handlers, but the framework must also provide mechanisms for handling possible problems or fill possible gaps addressed there. In the current state of the PERSONA system, however, multimodality is supported only by multimodal I/O handlers.

The second severe issue concerns the configurability of the system. Currently, there is no simple solution for end users that can be intuitively used for configuring the system, e.g. in the case of introducing new rules to the Situation Reasoner. It may be acceptable at the deployment time to assume that such tasks will be done by service providing experts, but at least in case of later adjustments it is preferable that end users can directly take corrective actions. This means that there is much potential to work towards a sort of end user programming.

References

1. Balme, L., Demeure, A., et al.: CAMELEON-RT: A Software Architecture Reference Model for Distributed, Migratable, and Plastic User Interfaces. In: Markopoulos, P., Eggen, B., Aarts, E., Crowley, J.L. (eds.) EUSAI 2004. LNCS, vol. 3295, pp. 291–302. Springer, Heidelberg (2004)
2. Bouchet, J., Nigay, L., Ganille, T.: ICARE Software Components for Rapidly Developing Multimodal Interfaces. In: Proceedings of ICMI 2004, pp. 251–258. ACM, New York (2004)
3. Boyer, J.M.: XForms 1.0 (third edition). W3C Recommendation, W3C (2007), http://www.w3.org/TR/2007/REC-xforms-20071029/
4. Calvary, G., Coutaz, J., Thevenin, D.: Supporting Context Changes for Plastic User Interfaces: A Process and a Mechanism. In: Proceedings of IHM-HCI 2001. Springer Conference Series, pp. 349–363 (2001)
5. Calvary, G., Coutaz, J., et al.: Plasticity of User Interfaces: A Revised Reference Framework. In: Proceedings of TAMODIA 2002, pp. 127–134. INFOREC Publishing House Bucharest (2002)
6. Coutaz, J., Balme, L., et al.: An MDE-SOA Approach to Support Plastic User Interfaces in Ambient Spaces. In: Stephanidis, C. (ed.) UAHCI 2007 (Part II). LNCS, vol. 4555, pp. 63–72. Springer, Heidelberg (2007)
7. Duarte, C., Carriço, L.: A conceptual framework for developing adaptive multimodal applications. In: Proceedings of IUI 2006, pp. 132–139. ACM, New York (2006)
8. Fides-Valero, A., Freddi, M., Furfari, F., Tazari, M.R.: The PERSONA Framework for Supporting Context-Awareness in Open Distributed Systems. In: Aarts, E., Crowley, J.L., de Ruyter, B., Gerhäuser, H., Pflaum, A., Schmidt, J., Wichert, R. (eds.) AmI 2008. LNCS, vol. 5355, pp. 91–108. Springer, Heidelberg (2008)
9. Flippo, F., Krebs, A.M., Marsic, I.: A framework for rapid development of multimodal interfaces. In: Proceedings of ICMI 2003, pp. 109–116. ACM, New York (2003)
10. Ambient Intelligence: from vision to reality. ISTAG Report, European Commission – IST Advisory Group (2003), cordis.europa.eu/ist/istag-reports.htm
11. Ju, W., Lee, B.A., Klemmer, S.R.: Range: exploring implicit interaction through electronic whiteboard design. In: Proceedings of CSCW 2008, pp. 17–26. ACM, New York (2008)
12. Martin, D., Paolucci, M., et al.: Bringing Semantics to Web Services: The OWL-S Approach. In: Cardoso, J., Sheth, A.P. (eds.) SWSWPC 2004. LNCS, vol. 3387, pp. 26–42. Springer, Heidelberg (2005)
13. Schmidt, A.: Implicit Human Computer Interaction Through Context. Personal and Ubiquitous Computing 4(2/3) (2000)
14. Tazari, M.R.: Using Queries for Semantic-based Service Utilization. In: Proceedings of SMR 2009. CEUR Workshop Proceedings, vol. 525, p. 15 (2009)
15. Tazari, M.R., Furfari, F., Ramos, J.P.L., Ferro, E.: The PERSONA Service Platform for AAL Spaces. In: Handbook of Ambient Intelligence and Smart Environments, pp. 1171–1199. Springer, Heidelberg (2009) ISBN: 978-0-387-93807-3
16. Weakliam, J., Bertolotto, M., Wilson, D.: Implicit interaction profiling for recommending spatial content. In: Proceedings of GIS 2005, pp. 285–294. ACM, New York (2005)

Appendix: Additional Notes

Input by the user is normally captured with the help of devices like a microphone for capturing voice, a camera for capturing gesture, or other devices, such as a mouse, a keyboard, or a remote control that work in the context of a display. How I/O handlers bind the input channels provided by such devices is in the responsibility of their developers. For binding microphones and cameras, the corresponding I/O handlers in PERSONA are using a solution as sketched in figure 5. Similarly, service components that wrap loudspeakers provide voice output services on the service bus that are utilized by the corresponding I/O handlers. The benefit for I/O handlers is their independence from concrete devices and their distribution in the environment while enabling the multi-purpose sharing of these devices.

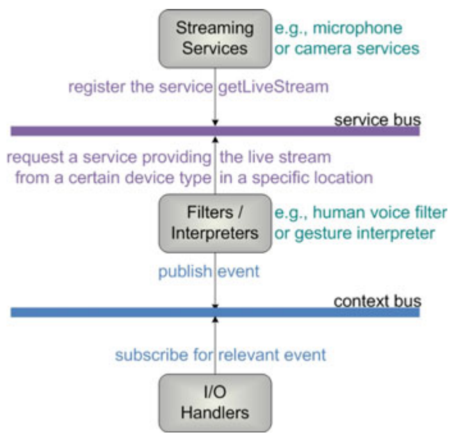


Fig. 5. Binding streaming devices as input channel

As of the rules for deriving the value of dynamic adaptation and profiling parameters, it is worth mentioning that the Situation Reasoner provides mechanisms for external configuration of the set of defined rules even at runtime. This is one of the bases that allows to keep the UI framework open in regard to user and context models.

A Vision-Based System for Object Identification and Information Retrieval in a Smart Home

Raphael Grech[2,⋆], Dorothy Monekosso[1],
Deon de Jager[2], and Paolo Remagnino[2]

[1] Computer Science Research Institute, University of Ulster, UK
[2] Digital Imaging Research Centre, Kingston University, UK

Abstract. This paper describes a hand held device developed to assist people to locate and retrieve information about objects in a home. The system developed is a standalone device to assist persons with memory impairments such as people suffering from Alzheimer's disease. A second application is object detection and localization for a mobile robot operating in an ambient assisted living environment. The device relies on computer vision techniques to locate a tagged object situated in the environment. The tag is a $2D$ color printed pattern with a detection range and a field of view such that the user may point from a distance of over 1 meter.

Keywords: Smart home, object detection and recognition.

1 Introduction

The work described in this paper is part of a larger project to develop assistive technologies to cater for people with physical and cognitive impairments. We are investigating methods to guide a person towards an object and *receive* information about the object from the object. The device described in this paper was conceived with one application in mind, although it can be adapted for other applications. The first application is a device to assist a person suffering from memory and other cognitive impairments (such as dementia, Alzheimer's disease, etc $\cdots$). The second application is an object detection and localization system for a mobile robot. In the first application, the device is standalone and contains a database of known objects in the environment. The feedback, information about the object, from the device is delivered as audio. In the second application, the device will assist and guide a robot in localizing an object. The robot operates as an assistant/service robot in a smart home. The device relies on computer vision techniques to locate a tag on an object when it is within the field of view of the device. The tag is conceptually similar to 2-dimensional labels that hold item or product information such as the data matrix or the Microsoft tag. However there are two main differences: the detection range and robustness

⋆ The authors wish to acknowledge the support provided by the Computer Science Research Institute at the University of Ulster.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 239–247, 2010.

to tremor/jitter. The target users are elderly persons, particularly those with memory and/or physical impairment. Robustness will ensure tremor or effects of an unsteady hand are removed. To satisfy these requirements, the range and the field of view (FOV) are such that the user (or robot) may point from a distance of over one meter. The tag will be detected as long as the FOV is within $\pm 30°$ and audio feedback will guide the user or robot towards the object until the user or robot faces the object. A small foot print for the tag, $57mm \times 70mm$, will ensure it can be placed on smaller household items. The detection range is constrained by the camera resolution and the tag size. The range and FOV will depend on the camera characteristics, but can be increased by using a higher resolution camera, or increasing the tag size. Information about the object pointed to is contained within a database held on the hand-held device. It is envisaged that a dedicated standalone device is developed, however the systems software may operate on a generic PDA or mobile phone with a camera. The remainder of this paper is organized as follows. A review of labeling technology and data processing techniques is discussed in Section 2. In Section 3, the methodology is described, followed by the experiments and results in Sections 4. We finally conclude our findings in Section 5.

2 Related Work

In this section, we begin the review with systems designed to assist with the task of finding and retrieving objects. We are particularly interested in systems that employ labeling technology to tag the object as this is the basis for the system described in this paper. First we consider two systems that use the existing characteristics of the object and do not tag. Merler et al in [7] describe a vision-based system to recognize groceries in situ using in vitro training data. The object recognition system is trained using image data captured under laboratory conditions (in vitro). Object recognition, in this case food items, then takes place, in situ, in a grocery store. The challenge is to recognize food items such as a can of soup in a visually cluttered environment. Narasimhan et al in [8] simplify the problem by reading barcodes found on most grocery items. The system, an aid for visually impaired shoppers, is based on a Smartphone. The types of labels catered for are UPC bar codes and RFID as found on items of clothing. The reading is at close range, in the range of millimeters, and the user must be guided towards the items. The two systems described above rely on reading the label and retrieving the information provided by the manufacturers of the labeled product. The systems described next do not use the manufacturer label but rather a proprietary labeling and information content system. There are a number of RFID-based localization and navigation systems found in the literature. We consider only a few here. Hallberg et al [4] localize household object which have been RFID tagged to assist persons with dementia; the localization algorithms employ Received Signal Strength Intensity (RSSI) on the reader. Kulyukin et al [6,5] also employ RFID tags for an indoor assisted navigation system; the RFID tags are dispersed in the environment and the reader located on a mobile robot platform. Moving away from RFID technology, we consider localization systems that use a 2D visual pattern. Labeling technology and label reading technology has evolved from the simple one-dimensional bar code read (scanned) by an optical decoder to two dimensional matrices known as a data matrix. The technology employs a machine-readable data code that is printed onto a tag which is attached to an object. There are a number of proprietary and open-source systems available; differing in the representation and encoding of the data which in turn determines the method used to read the code. The Microsoft Tag is mentioned here as it uses color to increase data capacity. The High Capacity Color Barcode (HCCB) uses clusters of colored triangles. The data density is increased by using a palette of 4 or 8 colors for the triangles. In a Microsoft Tag, the HCCB contains a hyperlink, which when read, sends the HCCB data to a Microsoft server, which in turn returns the publisher's intended URL. The Tag reader directs the user's mobile browser to the appropriate website. An assistive device employing machine vision, was described by Coughlan et al [2,3]. The focus is navigation for the visually impaired; employing a mobile phone camera to detect the color targets situated in the environment. The use of colored targets with specific patterns allows the target to be detected at a distance and in seconds.

3 Methodology

The system comprises three components: the label or tag, the label reader and information store. These three components are discussed next.

The Tag

The prototype tag is shown in Figure 1. It comprises five circles (dots); each circle can take on any of the 6 colors - red, green, blue, magenta, cyan, and yellow. The border, surrounding the circles, allows the tag to be detected, akin to the finder pattern in a data matrix and facilitates localizing. Orientation is determined with a marker indicating a point of reference. In our scenario, the color order is important and repetitions are allowed. This provides for n^r permutations, where n is the number of used colours and r the number of used dots. In our case,

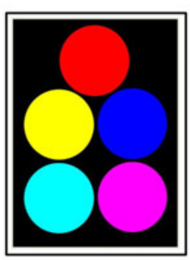


Fig. 1. Proposed Prototype Tag

we have $6^5 = 7776$ permutations. If repetitions are not allowed then we can cut down the number to 720 permutations according to

$$\frac{n!}{(n-r)!} \tag{1}$$

In the context of object recognition for ambient assisted living, this provides a sufficient number of unique patterns to identify all required objects. The rationale for the tags design, from a computer vision perspective, is ease of production, using any color laser printer; the use of primary colors facilitates segmentation, and orientation. The requirements are for the tag to be detectable from at least one meter with a 320×240 pixels camera and up to two meters with a 640×480 pixels camera (the target is detected up to $1/100^{th}$ of image size). The detection of the tag pattern in an upright position allows for a tilt/skew of up to $\pm 30°$. Detection must be robust under different lighting conditions and in a visually cluttered environment.

The Tag Reader

The data contained within the tag is extracted with a camera mounted on a small hand held device such as a smart phone so that standard computer vision techniques can be employed. Tag data extraction is performed in two steps: the detection of the overall tag shape followed by a segmentation and extraction of the colored circles.

Haar Classifier and Color Segmentation

The tag is detected when the finder pattern is recognized. The *Haar* classifier is employed to detect the finder pattern; it is trained in situ to recognize the finder pattern in a visually cluttered environment. The prototype uses the OpenCV library's implementation of the *Haar* classifier method. Details of this classifier can be found in [1]. The training set comprises only six of the possible 7776 patterns. The rationale for this choice was to investigate the performance with a low training set. Two training sets were created using different cameras for cross validation. The first dataset, set I, comprised images generated using two different hand held cameras and a robot camera. The second dataset, set II, was generated using the robot camera. During evaluation, validation was performed with dataset II when training was carried out with dataset I and vice versa. While the first dataset (set I) consisted of images of tags located randomly in the environment; dataset II was generated with the robot camera located at the node positions in Figure 2. The test image was captured at 30° interval. Once the tag shape is detected and recognized successfully, color segmentation is carried out based on the method proposed by Zhang et al. [9].

Information Store

The database contains information pertaining to the object recognized. The detected pattern points to a database entry.

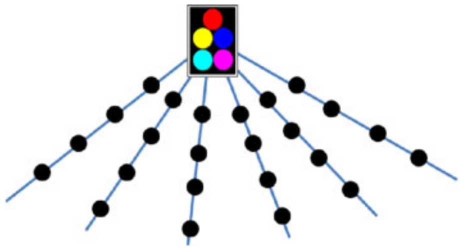


Fig. 2. Test pattern generation

4 Experimental Set-Up and Results

Software development was carried out using the OpenCV library [1] on a laptop connected to a robot and to a USB web camera. OpenCV is an open source development library written in C, and it is specifically suited for real time applications. Tests were carried out to investigate the praticality of our system. These tests involved: a) varying the distance and the orientation to determine the maximum detection range of the system, and (b) employing different cameras and light conditions to assess robustness and (c) exposing the system to severe jitter to reflect the condition of the actual user for whom this application is intended. Additional tests (d) were carried out to investigate performance with the tag applied to non-flat surfaces. In the experimental setup, the tag size was $57mm \times 70mm$ and the camera was set up to a resolution of 320×240 pixels. In practice, the system is able to detect tags as low as 10pixels $\times$ 10pixels; the limitation is to allow for color segmentation/detection within the small white boxes. Figure 3(a) and (b) show results with four tags. In Figure 3(a), the blue and cyan are adjacent in the color space however these were successfully identified, as shown in Figure 3(b). Tag detection capability of segments in the images affected by skew and affinity can be seen in both Figure 3(a) and (b). In the prototype, the detection area is fixed and contained within the small white boxes thus restricting the angle of detection to $\pm 30°$. This can be improved by employing a rotation detection algorithm. The top right tag and bottom left show

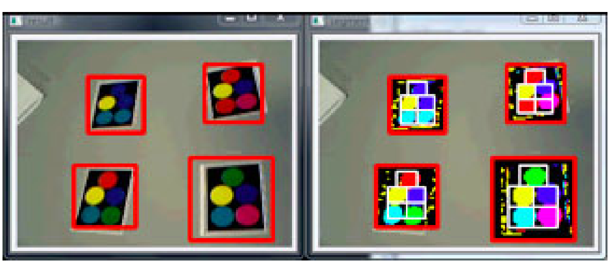


Fig. 3. Tag detection, image (a) and detection (b)

the current limitation; if the angle were to be increased further. The top right tag and bottom left show the limiting skew.

Detection Range and Orientation

During the testing phase it was expected that all the 6 patterns on which the training was carried out would be easily recognizable. A further 6 images comprised the independent test set. The independent test data were successfully recognized. Figure 4(a) and (b) show the detection percentage at a given distance and orientation respectively. The tests were carried out using the robot placed at an fixed distance and angle from the tag to be detected. Increasing detection range can be achieved with a higher resolution camera. It is also worth noting that for a 640 × 480 pixels resolution detection was achieved in real time. However the 320 × 240 pixels resolution, was used to enable the system to be integrated into of-the-shelf embedded systems with lower camera resolution.

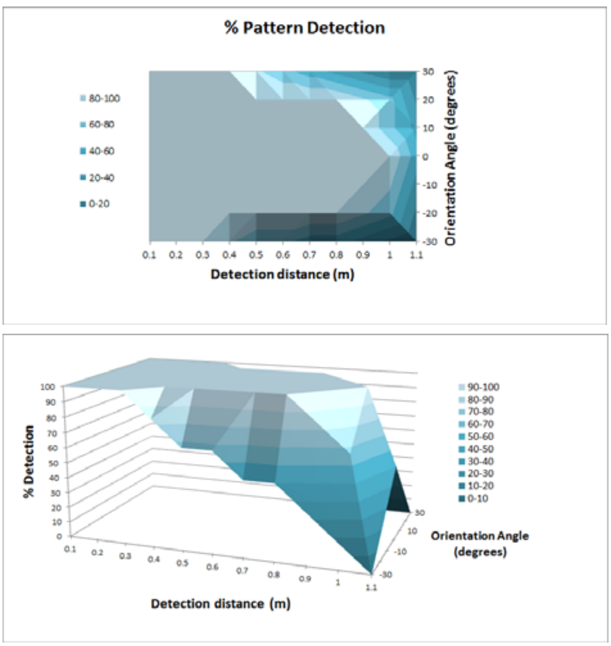


Fig. 4. Detection Range (a) and Orientation (b)

Detection under Different Light Conditions

The light intensity was varied from good lighting to dim lighting and the pattern was again successfully recognized and color segmented as shown in Figure 5.

Fig. 5. Different Light Intensities - (a) high and (b) low intensity

Detection under Conditions Simulating Tremor (Jitter)

As an assistive device, detection performance must be maintained when subject to tremors; as when used by a person suffering from Parkinson's disease or any other condition in which the user is unable to hold the camera steady. A 3D accelerometer was attached to a webcam to measure and coarsely classify the simulated tremor. The experiments were repeated with jitter included to simulate an unsteady hand. The results are shown in Figure 6, on the left (Figure 6(a)) clearly shows the blurring and the one on the right (Figure 6 (b)) is the real time detection of the patterns. The tags were successfully detected with various degrees of jitter. The detection limitation comes from blurring caused when velocities are high. It can be seen that system performs well under severe jitter.

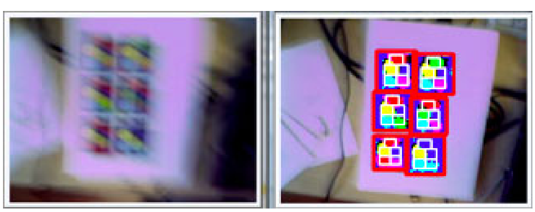


Fig. 6. Detection under conditions of jitter, (a) blurred image and (b) detected tags

Detection on Curved Surface

In a home environment, the tag may be fixed to a number of objects; some object might have a curved shape. When the pattern is fixed to a curved surface the circles will become ellipses, however, the *Haar* classifier, will detect these

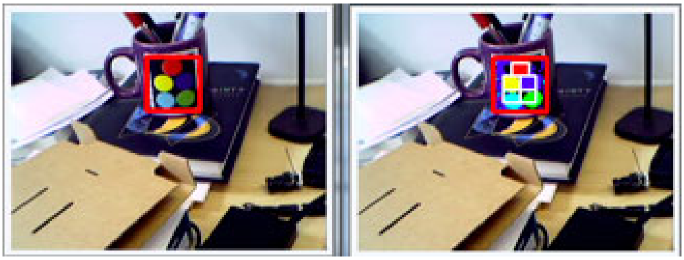


Fig. 7. Tag on a curved object

patters, even though training was carried out on flat surfaces. The result can be seen in Figure 7.

5 Future Work and Conclusions

Results have shown that tagged objects can be detected at over one meter in poor lighting conditions as well as under conditions of jitter. The efficiency of the device can be extended by increasing the range and robustness, particularly when there is significant visual clutter. We propose to employ error detection and correction in order to increase detection under conditions of poor indoor lighting condition and visual clutter. Different methods to convey the information about the tagged objects may be investigated. One possibility is to convert from text to speech and provide the information as audio.

References

1. Bradski, G., Kaehler, A.: Learning OpenCV: Computer Vision With The OpenCV Library. O'Reilly Media, Sebastopol (2008)
2. Coughlan, J., Manduchi, R.: Color targets: fiducials to help visually impaired people find their way by camera phone. EURASIP Journal on Image and Video Processing 2007(96357), 1–13 (2007)
3. Coughlan, J., Manduchi, R.: Functional assessment of a camera phone-based wayfing system operated by blind and visually impaired users. International Journal on Artificial Intelligence Tools 18(3), 379–397 (2009)
4. Hallberg, J., Nugent, C.D., Davies, R.J., Donnelly, M.: Localisation of forgotten items using rfid technology. In: International Conference on Information Technology and Applications in Biomedicine, pp. 1–4 (2009)
5. Kulyukin, V., Gharpure, C., Nicholson, J.: Robocart: Toward robot-assisted navigation of grocery stores by the visually impaired. In: IEEE International Conference on Intelligent Robots and Systems, pp. 2845–2850 (2005)
6. Kulyukin, V., Gharpure, C., Sute, P., De Graw, N., Nicholson, J.: A robotic wayfinding system for the visually impaired. In: International Conference on Innovative Applications of Artificial Intelligence, pp. 864–869 (2004)

7. Merler, M., Galleguillos, C., Belongie, S.: Recognizing groceries in situ using in vitro training data. In: IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2007)
8. Narasimhan, P., Gandhi, R., Rossi, D.: Smartphone-based assistive technologies for the blind. In: International Conference on Compilers, Architecture, and Synthesis for Embedded Systems, pp. 223–232 (2009)
9. Zhang, J., Zhang, Q., Hu, J.: Rgb color centroids segmentation (ccs) for face detection. International Journal on Graphics, Vision and Image Processing 9(II), 1–9 (2009)

SeSaMoNet 2.0: Improving a Navigation System for Visually Impaired People

Ugo Biader Ceipidor[1], Carlo Maria Medaglia[1], and Eliseo Sciarretta[1,2]

[1] C.A.T.T.I.D. – Sapienza University of Rome, Piazzale Aldo Moro 5, 00185 Rome, Italy
{ugo.biader, carlomaria.medaglia}@uniroma1.it, sciarretta@cattid.uniroma1.it
[2] Scuola IaD - Tor Vergata University of Rome, PhD program in Archiviazione e Gestione dei Documenti Digitali, Via della Ricerca Scientifica, 1, Rome

Abstract. The authors present the improvements obtained during the work done for the last installation of SeSaMoNet, a navigation system for blind people. First the mobility issues of visually impaired people are shown together with strategies to solve them. Then an overview of the system and of its main elements is given. Afterward, the reasons which brought to a re-design are explained and finally the main features of the last system revision for the application are presented and compared to the previous one.

Keywords: accessibility, navigation, RFId, blind, visually impaired, mobility.

1 Introduction

Independent mobility is recognized as a right for disabled people by the United Nations General Assembly, which stated in "The Standard Rules for the Equalization of Opportunities of Persons with Disabilities" [1] that "states should initiate measures to remove the obstacles to participation in the physical environment".

Still, for visually impaired people, pedestrian mobility is very difficult. To explore the space, sightless people often have to develop other senses (i.e. hearing, touch and smell) to compensate their disability, in order to create mental maps of places based on sensorial images. Moreover, they can rely on traditional assistive tools, as the white cane or the guide dog, to get help in reaching their destination without harm. However, this approach makes sense only if the person already knows, at least roughly, the path to the destination. Usually, the perception of a blind person only approximately extends to the tip of his/her cane, and so orientation is very difficult. Above all if environments are not designed for blind people.

In the last years, many researchers have been conducting studies aimed to solve the problem of independent mobility and to allow blind people to move relying only on modern technologies, without the involvement of other persons (or dogs).

The so called Assistive Technologies are aids created to support people with disabilities, in order to let them live independent lives in any facet [2].

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 248–253, 2010.

Many attempts have been done (e.g. in [3], [4] and [5]), but the real current challenge is to integrate new useful features in systems already used by blind people.

2 SeSaMoNet 1.0

SeSaMoNet, acronym for Secure and Safe Mobility Network [6], is a research project that aims to develop a reliable system to support independent mobility of visually impaired citizens in several settings, such as urban, cultural and naturalistic.

The SeSaMoNet solution couples tactile perceptions with hearing aids. To this extent, wireless technologies (RFId and Bluetooth), hand-held devices (Smartphones or PDAs) and specific system and application software for mobile device (e.g. Text To Speech, database, etc) are combined together.

SeSaMoNet is a European Commission project, developed by the Institute for Protection and security of the Citizen (IPSC) of Joint Research Centre in co-operation with CATTID labs of the "Sapienza" University of Rome and Unione Italiana Ciechi (UIC), the Italian blinds' association.

A great attention has been paid to the effectiveness and the usability of the system by the integration of innovative ICT's hardware and software components with the traditional aids. The ease of use and the confidence provided by the white cane are combined with the features supplied by new technologies.

2.1 Elements and System Description

The system is composed by the following elements:

- The RFId tag grid: the 134.2 KHz channel is used since radio frequency noise from interposing environment (i.e. water) is very reduced with respect to the signal.
- The RFId reader cane is designed to be as similar as possible to the one used by blind people. An RFId antenna is integrated in the tip of the cane, while the handle hosts the rechargeable batteries as well as the Bluetooth terminal.
- The smartphone (and Bluetooth earpiece) containing
  - the Text-To-Speech (TTS) software: the Nuance Talks screen reader is currently used;
  - the management system/navigation data server: the software has been developed in C/C++ under S60 3rd edition platform to be used on devices equipped with Symbian OS version 9;
  - the database, downloaded, for local information, from a server, when possible to be connected by WiFi.

The special cane reads the tag's unique ID by RFId communication. It then sends this information through a Bluetooth channel to the smartphone device. The ID string is needed as key to query the database; navigation data retrieved are processed, converted in auditory output through the TTS software and then sent through another Bluetooth channel to the earpiece.

The user can be given hints on whether he is going along the right path or not. He/she also can receive descriptions of nearby points of interest, depending on the setting of use. In fact, the system is suited for both indoor and outdoor.

2.2 Software Features

The main features that SeSaMoNet offers are the following:

- To keep the user inside a safe path;
- To provide information about turns and obstacles on the path;
- To check whether the direction is right (useful if user falls down or is disoriented);
- To provide specific environmental descriptions.

3 SeSaMoNet 2.0

The system as described above has been tested several times in many installations, either temporary (as demonstrations in fairs) and permanent (in museums across Italy and Europe), with excellent responses by the users, retrieved by means of questionnaires and personal interviews, where they remarked the usefulness of the service to achieve their own complete social inclusion; still something was missing.

Sesamonet project aims to be a valuable support for the independent mobility and living of visually impaired persons, integrated with traditional aids.

So the natural continuation of the project has been its integration with tactile paving systems, already broadly used to realize paths for blind people, in order to maximize the benefits of both systems.

The first chance has occurred during the installation of SeSaMoNet at the “Agostino Gemelli” hospital in Rome, wanted by the UIC. In this case, SeSaMoNet has been blended with the “Vettore” [7] paving system, produced by Antonplast [8] and supported by the UIC itself.

Both the hardware and the software of SeSaMoNet have been partially re-designed to fit the needs of the new setting.

3.1 Re-design

The hardware changes refer above all to the tag grid. In the previous version, each row of the grid was made up by three tags. Each tag was associated with a position along the path (right, left or centre) and each position was associated with a different short sound conveyed to the user. This configuration was necessary to create a safe corridor, in order to avoid letting the user out of the path.

In the current SeSaMoNet, instead, this task is accomplished by the path itself, which gives to the user tactile hints: when the user feels little bumps under the tip of the cane, he/she is sure to be on the right track. So there is no more the need for three tags per row; just one, placed in the centre, is enough to convey the key information.

The decrease of the number of needed tags makes possible (economically) the adoption of more efficient tags, above all in indoor settings. In the first version of the system, in fact, small flat disk-shaped tags were used (only 3 cm in diameter) to guarantee the non-interference between close tags. But these tags allowed only a short reading range (up to 10 cm), causing some skips in the reading.

Now the system can exploit larger tags (15 cm in diameter) placed one after the other at a distance of almost 100 cm, because they grant a much wider reading range.

In cultural environments, such as museums, where there is no way to put the tags directly in the floor, the tags were used to be placed under a runner-carpet rolled out the floor. This carpet, though, needed to be fixed to the ground to avoid accidental hindrances and this resulted in a too high impact in the environment.

So, following the example of tactile pavements, we designed a footbridge made of see-through methacrylate (Fig. 1). The footbridge is composed by modular elements, which combined together weigh enough to stay still, so they can be easily leaned on the floor, without any fixing. The small vertical height of the footbridge (only 1 cm of ground elevation) and its beveled edges prevent the risk of accidental stumbling.

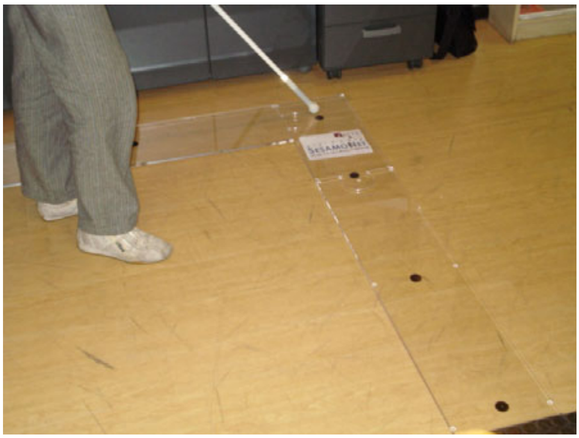


Fig. 1. The footbridge used for SeSaMoNet 2.0

3.2 New Features

The main innovative changes were applied to the management system, in which many new features were added.

First of all, the user has now the chance of choosing how many tags the cane has to read: each tag (the application bleeps every time the cane reads a tag), none (no blip is uttered, only the descriptions about the environment are conveyed to the user) or once in a while (i.e. the user can choose to have a blip every five tags read, this option is useful to be certain that the system is working). This feature is accessible through an options menu when the application is launched.

More important, a new system for the management of nodes has been developed: when the user approaches a crossing, the application can dynamically compute the choices at user's disposal, just knowing the segment he/she is coming from. The user listens to the possible destinations and he/she can start walking in any direction: the application will recognize the undertaken path and communicate it to the user as security feedback. This way, the user has the most freedom of choice, obviously always remaining above the arranged footbridge.

A redundancy control has also been added. Often the same piece of information is inserted in two or more consecutive tags in order to assure that specific indication is conveyed to the user. The application is able to recognize whether a tag containing the same indication has been already read and in that case it isn't repeated anymore.

A great attention was paid to the accessibility of the system: now the application comes in form of a single auto-installing packet, which also contains the database. Bluetooth pairing between cane and device has been simplified since the user has only to choose the right device from a list of available devices detected by the phone (through the same options menu described above).

4 Conclusions and Future Work

SeSaMoNet 2.0 mixes the well-established features of the first version with a real integration of assistive technologies, in order to simplify independent mobility of blind people. The system composed by SeSaMoNet and the tactile pavements is embedded in the urban environment, together with audible traffic signals, in the perspective of realizing the concept of "smart cities" [9].

In a designed-for-all city, where all architectural barriers are removed, SeSaMoNet becomes an integral part and can provide additional intelligence by supplying further information, granting at the same time safety of movement to blind people.

The next steps in the SeSaMoNet project will be a massive test session with users at the Gemelli installation to evaluate the efficiency of the integration between the system and the Vettore pavement, in order to proceed in further development and the spread of the solution with installations across Italy or anywhere it will be desired.

Acknowledgment

This study has been supported by the Ministry of Regional Affair Initiative PROGRAMMA ELISA – WI-MOVE project and by EUMETSAT study contract EUM/CO/04/1386/KJG. The authors wish to thank UNIONE ITALIANA CIECHI and ANTONPLAST srl for his precious inputs.

The authors wish also to thanks the SESAMONET team of the JRC – ISPRA (and in particular Dr. Graziano Azzalin) for the useful conversations on this topic.

References

1. The Standard Rules for the Equalization of Opportunities of Persons with Disabilities, Adopted by the United Nations General Assembly, forty-eighth session, resolution 48/96, annex (December 20, 1993), http://www.un.org/esa/socdev/enable/dissre00.htm
2. Morini, A., Scotti, F.: Assistive Technology. Tecnologie di supporto per una vita indipendente. Ed. Maggioli (2005)
3. Cardin, S., Thalmann, D., Vexo, F.: Wearable System for Mobility Improvement of Visually Impaired People. In: VR Workshop on Haptic and Tactile Perception of Deformable Objects, Hannover, Germany (December 2005)
4. Kobayashi, Y., Takashima, T., Hayashi, T., Fujimoto, H.: Gait analysis of people walking on tactile ground surface indicators. IEEE Transaction on Rehabilitation Engineering (2005)

5. Kulyukin, V., Gharpure, C., Nicholson, J., Pavithran, S.: RFID in Robot-Assisted Indoor Navigation for the Visually Impaired. In: Proc. of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004), Sendai, Japan (2004)
6. Biader Ceipidor, U., Medaglia, C.M., Azzalin, G., et al.: A RFID System to Help Visually Impaired People in Mobility. In: EU RFID Forum 2007, Brussels, Belgium, March 13-14 (2007)
7. `http://www.sinusoide.it/antonplast_web.pdf`
8. `http://www.antonplast.it`
9. Komninos, N.: Intelligent Cities: Innovation, knowledge systems and digital spaces. Taylor and Francis, Spon Press, London, New York (2002)

Plugin Driven Architecture for Intelligent Management of Building

Alessandro Olivi, Roberto Borsini, and Alessandro Bastari

Gruppo Loccioni - Via Fiume 16, 60030 - Angeli di Rosora (AN), Italy
{a.olivi,r.borsini,a.bastari}@loccioni.com
http://www.loccioni.com, http://www.leafcomunity.com

Abstract. In the present work an innovative solution based on a middleware framework is introduced which simplifies the modeling of a software for intelligent management of buildings and allows a multistandard and multiprotocol integration of sensors and actuators. The adopted programming methodology has never been used in the implementation of a BA system and makes it possible to overcome some limitations of traditional approaches, such as the hard integration of new interfaces and protocols. The results of the framework application in a residential building proves the validity of the proposed solution.

Keywords: Model Driven Architecture, Building Automation.

1 Introduction

Within the last century a huge technological evolution occurred in both private and public buildings, as well as in the automotive sector but with remarkable differences. Nowadays, in fact, car manufacturers tend to neither produce nor realize every single component of their products but rigorously define their standards to guarantee a seamless integration; every vehicle has also a coordination power unit of the subsystems, also used to make a diagnosis in case of damage. These precautions also led to improve the comfort, to optimize consumptions, to repeat the model production, to decrease the planning timing and consequently the vehicle sale price. Despite the advantages, the same measures are not taken in buildings planning and even if in modern buildings there are numerous autonomous subsystems (boilers, anti-theft devices, solar panels, domotic systems, etc.), they generally operate independently and often custom objects are realized by manufacturers not respecting either physical or software standards.

The cooperation of the different subsystems of a building, in terms of sharing information and activities coordination is really fundamental for optimising the functionalities and integrating predictive logics. This aspect becomes still more important in a *smart grid* vision [1] where every building cooperates with other surrounding buildings to exchange energy and information on the relative productivities and needs. Modern buildings must therefore be conceived as "intelligent" and continuously evolving entities where to require new functionalities or replace the old ones. The commercial *Building Automation* (BA) frameworks

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 254–258, 2010.

currently available are proprietary, closed and allow the integration of a limited number of standards. In such a way, they do not provide an abstraction level oriented to scalability and interoperability, so to make technology mixing and later expandability of the system very cumbersome.

A first fundamental step towards the concept of *modern building* consists in the realization of a system of BA able to interface and coordinate all the subsystems found in the building for the functionalities and services technological evolution and the simplification of the new interface and protocols integration. In this work an innovative BA control system is introduced, composed of a scalable and adaptive software framework which allows the simple integration of new standards and protocols. The software architecture of the system is described in Section 2. The flexibility and optimisation capabilities of the framework are checked in a residential building, the *Leaf House*- see the Section 3.

2 Framework Architecture

The main characteristic of the solution proposed for the BA system is to implement a wrapper for each subsystem, named *driver*. The coordination platform consists in a framework [2] that carries out interoperability among different drivers. This middleware application implements the communication interfaces and makes the abstraction of subsystems possible. The development of a framework must aim at realizing a software platform which guarantees the fundamental requirements of Reusability, Modularity and Expandability (RME): **Reusability** means that a framework must include all the elements used in different situations, so to not duplicate the implementation of the same functionalities; **Modularity** implies that a framework includes the set of basic functionalities for the development of new components; **Expandability** to create a framework which can be easily expanded by introducing new functionalities with no need of modifying the application platform architecture.

Considering what said above, the Model Driven Architecture (MDA) [3,4] paradigm has been employed for software implementation. This software design methodology allows to split modeling from the realization of applicative components. The main advantage of the chosen strategy lies on the separation between model and technology. The BA framework can be seen as a software gateway to simplify the integration between subsystems realized with heterogeneous standards. Each interface component of domotic standards, each element of the Graphical User Interface (GUI) and each communication service is identified as a *plugin*.

The framework plugin architecture allows a seamless integration of new hardware by just implementing a new access driver to the physical device. This methodology makes a quick integration of new functionalities possible and does not require any modification to already tested and correctly running software. By following the MDA programming principles, a software development procedure based on Unified Modeling Language (UML) [5] and on the separation of applicative components in different abstraction levels has been used (Figure 1).

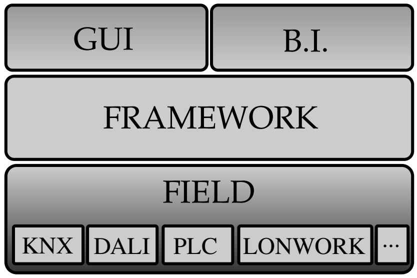


Fig. 1. Middleware Architecture

The field layer can be found at the lowest logic level. The layer is composed of all the drivers managing the standards interfacing. The framework layer is the central part of the application infrastructure and it is where the functioning logics are defined. The framework ensures that each plugin respects the interfaces defined during the design phase and provides the cooperation mechanisms among components. It also takes charge of the dynamic loading of plugins and services. The most significative functionality of the framework layer is the *virtual channel*.

The channel is a logic tool which makes the measure of the physical phenomenon, coming from a driver, independent from the information. The information can be normalized, analyzed, aggregated by following different logic and mathematical procedures and finally stored. Following processings can induce actions which alter the system status. The highest logic level contains the GUI and Business Intelligence (BI) analysis systems. The GUI can be remoted and adapted to the kind of display device used. The highest layer of the stack is independent from the specific framework instance so that the graphical components can gather information from different standards and from network distributed acquisition units. The channels sampled data recover measures that are stored in log files, databases or similar. By using a BI system, it is possible to develop analysis models depending on the specific application. The final user can analyze processed data through web pages or by using graphical plugins linked to the GUI level of the framework.

Following the MDA methodology, the Platform Independent Model (PIM) has been designed making use of different Design Patterns . The Platform Specific Model (PSM) has been realized by adapting the business function model to Microsoft.Net© software platform.

The most innovative characteristics of the framework developed can be summarized in the following points.

1. **Plugins dynamic loading:** the system must not stop and it is necessary that new plugins are dynamically added with a simple configuration;
2. **Communication Fault Tolerance:** for medium dimensions buildings the control system is distributed on more nodes with communication on IP backbone; in case of communication disservices, dedicated managers signal the disservice and start a mechanism for managing queues and transaction logs

which store the sampled information in a local cache. When resetting the connection the communication service will deal with data synchronization.

3. **Virtual Channels** is the main architecture element implementing the whole datum measurement cycle and its transformation into information; it is an abstract concept with which to correlate information coming from heterogeneous standards. The channels may be aggregated according to mathematical logics such as validity range, minimum, maximum, average in time and with the possibility to carry out analysis on the datum, algorithms injection realized with a high level programming language. The channel can be configured respecting the functional specifications of the application to be realized.
4. **Events management:** an event mechanism has been introduced for the activation of the framework functionalities only when the value measured by a sensor changes , when a threshold value is exceeded or an alarm is released so to decrease the calculation resources used. The channels can be configured both for polling or for event - driven data management. From the experience matured the polling is usually applied to a minimum part of the measures.

The framework currently owns a software library including the access drivers to the most diffused standards used in the domotic field (like Konnex [6], LonWork and Dali), communication driver (Web service, GSM/GPRS), communication driver towards the main PLCs and drivers for the alarm units, weather stations, photovoltaic systems inverters interfacing. Another advantage is that the drivers can be realized in parallel by different people who, using the SDK, must not necessarily know the whole framework infrastructure.

3 Case Study: The Leaf House

The previously described BA system has been used for the management of a technologically innovative residential unit of six apartments, named Leaf House. In the Leaf House more than 1.000 sensors and actuators have been interfaced such as geothermal pump for the hot and cold generation necessary to climatisation, alarm systems, electric loads management, thermal solar system, photovoltaic system, UTA systems for air purging, weather stations, etc. The coordination of all technological systems has been realized through a control system distributed on seven nodes (seven embedded PCs), one per apartment plus one for the common services management. Every node is governed by a BA framework instance. The single node is physically interfaced with the local subsystems. The virtualized sampled measures are made public respecting the policies hiding the information subject to safety limitations of each building. The nodes cooperate exchanging data on the IP backbone. Every instance of the framework enters the local measures and the ones provided by the other nodes.

The overall structure makes the data exchange management among nodes flexible and does not impede any evolutions of the infrastructure. The use of different domotic standards on buses and the system distributed architecture has reduced the structural wirings and made the system scalable for future extensions or modifications. The graphical interface of the system is supplied through touch-screen displays dislocated in the apartments. The information managed by every application are recovered by one or more physical nodes of the network. The data monitored are stored and analyzed by using a BI system. Electrical and thermal energy production and consumption data are managed as well as some comfort and indoor air quality parameters , the level of humidity and the temperature. The BA framework has also a field simulation driver. The simulator allows to define the whole BA system installation before the building effective realization.

4 Conclusions

Adaptivity and scalability of the presented BA system were proved through the case study described in Section 3. First of all, the work methodology allowed the contemporaneity between the hardware installation and the control functional definition, so that just two weeks of work at the building were necessary for the system startup. The event-driven management of the field level allowed the use of low performance CPUs leading to savings in both purchasing and energy consumption. Moreover the continuous control of the Leaf House environmental parameters determined a high level of perceived comfort (humidity and controlled temperature, air quality, utilities control, etc).

The BA framework application in the energy management and the technological measures adopted led to a total annual emission reduction of CO_2 of 21.000 kg respect to equivalent traditional building consumptions[1].

In the future, this solution will be integrated in complex systems for the man-building and building-outside world interaction.

References

1. The Eurpean Electrcity Grid Initiative (EEGI) - Roadmap 2010-2018 and Detailed Implementation Plan 2010-2012, http://www.smartgrids.eu
2. Medvidovic, N., Taylor, R.N.: A Classification and Comparison Framework for Software Architecture Description Language. IEEE Trans, Softw. Eng. 26(1), 70–93 (2000)
3. Frankel, D.S.: Model Driven Architecture: Applying MDA to Enterprise Computing. John Wiley and Sons, NY (2003)
4. Poole, J.D.: Model-Driven Architecture: Vision, Standards and Emerging Technologies. In: ECOOP 2001, Workshop on Metamodeling and Adaptive Object Models (2001)
5. Fowler, M., Scott, K.: UML Distilled: a Brief Guide to the Standard Object Modeling Language, 2nd edn. Wesley, Boston (2000)
6. Sakellaris, J.K.: Supervisory control using EIB - KONNEX Technology: a Sensor Network Protocol Enabling a Holistic and Environmental Approach in Architecture. In: IASME/WSEAS 2008, USA (2008)
7. Carbon Trust Standard Organization, http://www.carbontruststandard.com

[1] The comparison is referred to official data relative to traditional house in equivalent climate zone. The amount of CO_2 saved is calculated by multiplying the energy saved by a conversion coefficient as stated in [7].

Enhancing the Expressiveness of Fingers: Multi-touch Ring Menus for Everyday Applications

Dietrich Kammer[1], Frank Lamack[2], and Rainer Groh[1]

[1] Technische Universität Dresden, Fakultät Informatik, Professur Mediengestaltung, D-01062 Dresden, Germany
[2] T-Systems Multimedia Solutions GmbH Riesaer Straße 5, D-01129 Dresden, Germany
dietrich.kammer@tu-dresden.de, frank.lamack@t-systems.com

Abstract. In the future, environments equipped with interaction surfaces sensitive to touch input will be a key factor to enable ambient intelligence. The user's fingers become input devices, allowing simple and intuitive interaction, like the manipulation of digital objects. However, people are used to everyday applications offering a wide variety of menu choices. We evaluate ring menus to enhance the expressiveness of finger interaction on multi-touch devices. Applicability and limits of ring menus are discussed with regard to our implementation by means of a preliminary user study.

Keywords: Multi-touch, ring menus, pie menus, user interface design.

1 Introduction

Current multi-touch applications hint at future interaction with ubiquitous, intelligent surfaces in the context of ambient intelligence (cp. Fig. 1). However, they are generally limited to manipulation gestures. In contrast, even simple desktop applications are filled with menus and a multitude of different options. The same is conceivable for services and functions delivered by ambient intelligence technology in general.

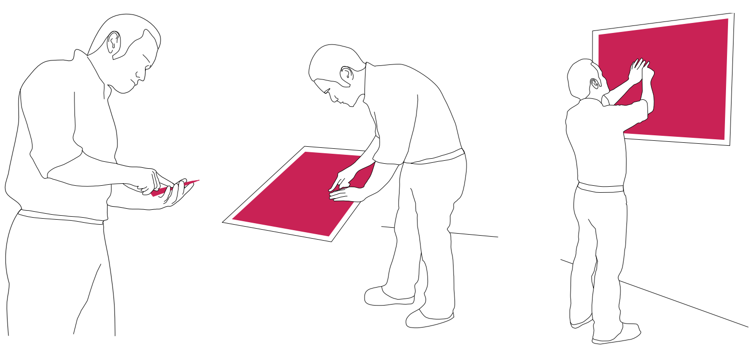


Fig. 1. Different layouts, sizes, and orientations of multi-touch surfaces

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 259–264, 2010.

The expressiveness of fingers can be enhanced by menu structures that are easy to use and adapted to multi-touch interaction, leveraging the potential of direct manual interaction, such as motor memory [8] and two-handed input [3,5].

We specifically address use cases from everyday applications like photo browsers or music libraries. This is due to their wide spread use, relatively low menu complexity, and compliance to imprecise user input. In contrast, applications for computer aided design or software developing environments require discrete devices for precise input, making them less accessible for ambient intelligence, which thrives on invisible technological means. Multi-touch technology meets this requirement by transparently transforming tables or walls into interactive surfaces operated directly without intermediary devices. In this contribution, we present related work concerning circular menu structures, an implementation of ring menus, and a preliminary user study which motivates future work.

2 Related Work

Ring or pie menus have been subject of research for many years in the HCI community. Performance benefits have been investigated and approved, e.g. in comparison to linear menu structures [6]. The circular layout shortens interaction paths, but limits the amount of items on each level of the menu hierarchy.

Pie menus allow both a textual and graphical representation of these items. Since the complexity of pie menus is low, they “permit low cognitive load actions” [11], which is beneficial to ambient intelligence applications that surround the user in his or her everyday activities. Similar research is centered on multi-touch marking menus, showing comparable benefits [9]. Pen-based interaction and issues of occluded menu items have been researched by [4].

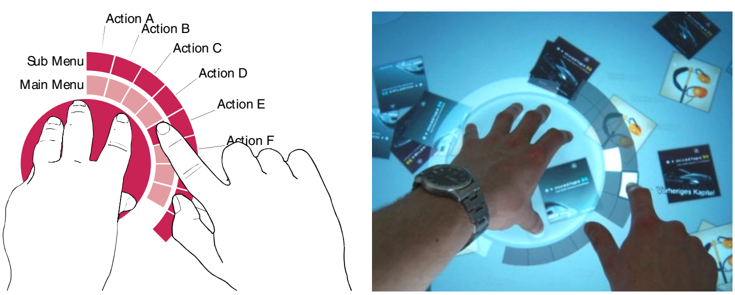


Fig. 2. Global menus substitute menu bars and can be activated using five fingers

3 Ring Menus and Multi-touch

Our hypothesis is that both the limited amount of menu items and short navigation paths are beneficial to multi-touch applications in the context of ambient intelligence. Similar observations are discussed on the web [7]. Especially in the case of large displays (cp. Fig. 1, middle and right), ring menus appear where users activate them, making them easier to reach in comparison to static menus. Their complexity is inherently kept low and motor memory can be leveraged, once an interaction path in the ring menu is learned [8]. Since we focus on multi-user interaction on large displays, two-handed input is possible. We propose global menus that are activated by the non-dominant hand, setting the size, rotation, and location of the menu. The dominant hand can select items and customize the rotation of the menu as well, using the left part of the ring, where no items are displayed (cp. Fig. 2). Context menus are activated by one finger that sets the orientation of the menu, avoiding occlusion of menu items (cp. Fig. 3). Subsequent selection is performed with the same or any other finger of both hands. We omit icons in ring menus which would require visual memory of users, necessitating correct interpretations [10]. Instead, text labels are assigned, which can later be ignored or removed by expert users trailing through the menu.

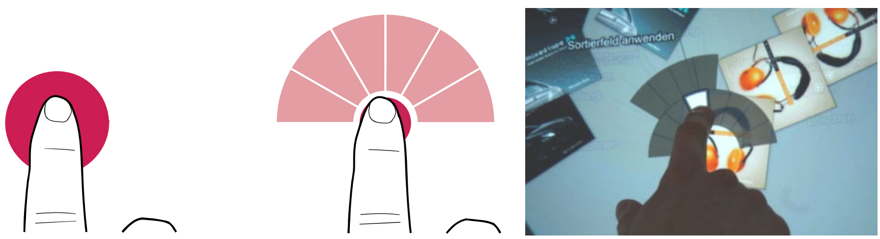


Fig. 3. Context menus are activated by a tap-and-hold gesture

4 Implementation

In order to test our concept for multi-touch ring menus, we implemented the *Radial Interface Toolkit (RITOK)* using the Microsoft® Surface SDK. Especially the provided finger orientation was crucial to this decision, since we aim to support table setups, which require menu orientation according to the user's position. RITOK offers reusable controls for global and context menus and includes events concerning the five fingers gesture. We applied RITOK to a music library with a simple provisional layout, substituting the menu-bar and context menus from popular desktop applications like Apple's iTunes by multi-touch ring menus (cp. Fig. 2).

5 Preliminary User Study

Our goal for the preliminary test of the prototype was to find errors, bugs, and user preferences for further development. Quantifiable information was not collected, but general feasibility of a user study concerning RITOK was approved. We interviewed 7 people with a technical background, all of whom had experience with touch-screens and multi-touch, but reported a low usage of these technologies. The questionnaire was structured according to ISO 9241-110:2006 (Ergonomics of human-system interaction, dialogue principles). In addition, four new criteria for multi-touch controls are introduced (dimensionality, spatial layout, ergonomics of hands, radial menus). The items of the questionnaire and conclusions are presented briefly below.

Suitability for the task. The dialogue should be suitable for the user's task and skill level. Test users agreed on the overall suitability of the menus to the task of organizing a music library.

Self-descriptiveness. The dialogue should make it clear what the user should do next, without any previous knowledge. The overall impression was that the structure of the ring menus is understandable. Additional descriptive information was judged insufficient and could be improved in the next iteration.

Controllability. The user should be able to control the pace and sequence of the interaction. The menu options could be executed fairly easy. Errors of the prototype were reported during recognition of the five fingers gesture.

Conformity with user expectations. The dialogue should be consistent in its behavior and it should meet expectations of users. The prototype proved to lack conformity, especially with regard to the activation time of context menus. Hence, a follow up study has to examine user expectations very carefully.

Error tolerance. The dialogue should be forgiving to false or unintentional input of the user. Error tolerance was weak due to the early state of the implementation.

Suitability for individualization. The dialogue should be able to be customized to suit the user. Individualization was judged moderately well.

Suitability for learning. The dialogue should support learning. The structure of the menus was found to support learning by the users involved in the preliminary study.

Dimensionality. This criterion was introduced to let users judge the efficient use of space of the menus on the multi-touch surface. An important factor is the reachability of target areas. Users reported that the concept allowed good or very good results with regard to this aspect.

Spatial layout. Important to the spatial layout is the adaptation of the ring menu to the user's position and interactional trajectories [2]. Regarding this criterion, only positive feedback was recorded.

Ergonomics of hands. Participants were asked to judge whether a comfortable posture of hands was achieved. This was negated, prompting improvements to the prototype and recognition of more comfortable hand postures (cp. Fig. 2). Another aspect is the size of target areas for activation of menu choices. Users agreed on the prototype offering an optimal size for fingers. The quality of the five fingers gesture to execute commands was inquired explicitly. Only moderate agreement could be ascertained which we subscribe to the uncomfortable posture forced on the hand by the prototype. Our initial approach to enforce bimanual interaction was tested and users expressed rather weak consent. Global menus should be controllable by a single hand as well.

Radial menus. The radial layout of the menus was judged very apt by the participants of our preliminary study. The theory that ring menus offer a great potential for multi-touch surfaces can thus be pursued further.

6 Discussion and Future Work

In this paper, we presented a working implementation of multi-touch ring menus. We matched potential strengths and limitations of multi-touch interaction to the layout of ring menus. For the first time, we present global ring menus that can be controlled with two-hands. Our preliminary user study shows that future work is needed in this field. An alternative to enhance the expressiveness of fingers over standard manipulation gestures are stroke shortcuts [1]. Future user studies have to compare differences and benefits of both approaches. Moreover, different strategies to menu activation and item selection have to be investigated. It is conceivable that ring menus actually rotate under a moving finger, further shortening interaction paths. Future work is needed to improve the orientation and placement of text labels, the stability of the prototype, and to set up a complete user study. The application of multi-touch technology and ring menus in the context of multimodal interaction is an ongoing research topic in the field of ambient intelligence.

Acknowledgements. We thank Ronny Meißner for his contributions to this research, graphics, and implementation efforts. Thanks are due to T-Systems Multimedia Solutions GmbH for providing the Microsoft® Surface and funding parts of this research. This work has been supported by the European Union and the Free State Saxony through the European Regional Development Fund (ERDF).

References

1. Appert, C., Zhai, S.: Using strokes as command shortcuts: cognitive benefits and toolkit support. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009, Boston, MA, USA, April 04-09, pp. 2289–2298. ACM, New York (2009)
2. Benford, S., Giannachi, G., Koleva, B., Rodden, T.: From interaction to trajectories: designing coherent journeys through user experiences. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009, Boston, MA, USA, April 04-09, pp. 709–718. ACM, New York (2009)
3. Bolt, R.A., Herranz, E.: Two-handed gesture in multi-modal natural dialog. In: Proceedings of the 5th Annual ACM Symposium on User Interface Software and Technology, UIST 1992, Monteray, California, United States, November 15-18, pp. 7–14. ACM, New York (1992)
4. Brandl, P., Leitner, J., Seifried, T., Haller, M., Doray, B., To, P.: Occlusion-aware menu design for digital tabletops. In: Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems, CHI 2009, Boston, MA, USA, April 04-09, pp. 3223–3228. ACM, New York (2009)
5. Buxton, W., Myers, B.: A study in two-handed input. In: Proceedings of CHI 1986, pp. 321–326 (1986)
6. Callahan, J., Hopkins, D., Weiser, M., et al. (Publisher): An empirical comparison of pie vs. linear menus. In: Proceedings of the SIGCHI Conference on Human factors Incomputing Systems, ACM, New York (1988)

7. Kiermasch, J.: The Comeback Of The Pie Menu. Centigrade Blog, June 16 (2010), http://www.centigrade.de/en/blog/article/the-comeback-of-the-pie-menu/
8. Krakauer, J.W., Shadmehr, R.: Consolidation of motor memory. Trends in Neurosciences 29, 58–64 (2006)
9. Lepinski, G.J., Grossman, T., Fitzmaurice, G.: The design and evaluation of multitouch marking menus. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems, CHI 2010, Atlanta, Georgia, USA, April 10-15, pp. 2233–2242. ACM, New York (2010)
10. Raskin, J.: The Humane Interface. In: New Directions for Designing Interactive Systems. Addison-Wesley Longman, Amsterdam (2000)
11. Shneiderman, B., Plaisant, C.: Designing the user interface - strategies for effective human-computer interaction. Pearson/Addison-Wesley, Boston (2005)

Privacy Policy Enforcement for Ambient Ubiquitous Services

Were Oyomno, Pekka Jäppinen, and Esa Kerttula

Lappeenranta University of Technology,
P.O. Box 55332, Lappeenranta, Finland
{were.oyomno,pekka.jappinen,esa.kerttula}@lut.fi
http://www.lut.fi/

Abstract. Ubiquitous service providers leverage miniaturised computing terminals equipped with wireless capabilities to avail new service models. These models are pivoted on personal and inexpensive terminals to customise services to individual preferences. Portability, small sizes and compact keyboards are few features popularising mobile terminals. Features enable storing and carrying of ever increasing proportions of personal data and ability to use them in service adaptations. Ubiquitous services automate deeper soliciting of personal data transparently without the need for user interactions. Transparent solicitations, acquisitions and handling of personal data legitimises privacy concerns regarding disclosures, retention and re-use of the data. This study presents a policy enforcement for ubiquitous services that safeguards handling of users personal data and monitors adherence to stipulated privacy policies. Enforcement structures towards usability and scalability are presented.

Keywords: Context-aware, privacy policy, ubiquitous service, privacy.

1 Introduction

The current generation of mobile terminals are miniaturised computing devices equipped with vast networking capabilities, built-in sensors and decent processors. Operating systems and applications executed on these processors facilitate activity scheduling/monitoring, contacts management and storage of personal data. Terminals compact form factor places them often in owners possession exposing them to identical environment as the host. Terminals capabilities coupled with hosts proximity exposes a mature platform to facilitate and enhance services to specific users' in specific situations.

Specificity requires services to adopt identities for distinguishing users and responding to individual situations/preferences. Ubiquitous services automate these requirements thereby raising privacy concerns among users. Two key concerns are centred on services credentials and their data handling practises. Data handling practises involve documenting personal data needed and adopted privacy measures regarding: safe-guards, sensitive data leaks, third party disclosures, privacy assurances, compromise compensations and prior privacy agreements. Similar privacy concerns will be unearthed and intensify as terminals further proliferate,

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 265–269, 2010.

sensing technology advances, more personalised services and the service industry matures. Unmitigated these concerns threaten ubiquitous services adoption.

Digital certificates are a viable solution to credential concerns while data handling concerns are partly addressed in web platforms with policy specifications. Common policy specifications are; privacy seals[1][2][3][4], Platform for Privacy Preference (P3P)[5], eXtensible Access Control Mark-up Language (XACML)[6], Security Assertion Mark-up Language (SAML)[7]. Lightweight cryptography in credentials validation have had more success than specifications in ubiquitous service environments. Major impediments to adoption of policy specification are their insufficiency to address users' questions regarding; what data are collected?, why are they collected?, how do data owners benefit?, how long are they retained?, who else see the data?, what is the service providers reputation?.

We address these policy specification impediments with an enforcement authority system. Our enforcer system verifies that specific services solicits user data within a privacy threshold. Solicited data are weighted against the privacy threshold in composing service policies. Each service has a policy (individual/group) known advance and register at the enforcement authority before being certified. Certified policies are the currency used in user-service interactions. Binding of enforcement authority, certifying authority, service identity and their privacy practises in their interactions serves to: (1) notify of data solicits, (2) scope retrieved data, use and purpose, (3) expose avenues for redress of non-compliancy and (4) limit malice by services.

2 Related Works

Efforts at user privacy in ambient environment often are modifications of Internet approaches based on: notice, processing time and cryptography. Notice approaches present users with standardised symbol confirming their adherence to basic privacy. TRUSTe and BBBOnLine privacy seals are common examples[1]. Processing time mechanisms specify and compare a user's privacy expectation against a service's data request during service consumption. E.g. P3P requires services to declare their intended use of personal data. The declaration comprises of collected data types, their permanence and visibility. XACML and SAML take a different approach. XACML standardises access control and authentication in manners describing entities and their attributes while, SAML is focuses on standardising the exchange this information between identity provider and service provider [1][3][6][7]. Cryptographic techniques to achieve anonymity, confidentially and integrity are usually integrated into processing time approaches. However, independently applied, they also provide anonymity (mixes, crowd, onions), confidentiality (Transport Layer Security (TLS), Secure Sockets Layer (SSL)) and integrity (message digest, signatures)[8][9].

Sealing a service does not guarantee it privacy preserving mechanisms. On the contrary, LaRose et al[4] notes that sealed services tended to request more privacy invasive information than unsealed services. Furthermore, seal providers make no commitments regarding the levels of privacy offered by services bearing their seals. Despite these limits, seals play a pivotal role in standardising how personal data requests are presented.

3 Privacy Policy Enforcement Architecture

Architecturally, the privacy policy enforcement system is composed of five key components: privacy policies (PP), personal trusted devices (PTD), service providers (SP), certifying authority (CA) and policy enforcement authority (EA). These components and their relations are depicted in Figure 1. Two service scenario, the information screen and the Ad-pusher are depicted. The information screen displays content personalised viewers (Alice, Bob) preferences while the Ad-pusher transmits preferred advertisement to users (Charlie, Dave) PTDs. Both service transact over Bluetooth access point (AP) due to; close proximity, lower cost and built-in service discovery protocol (SDP).

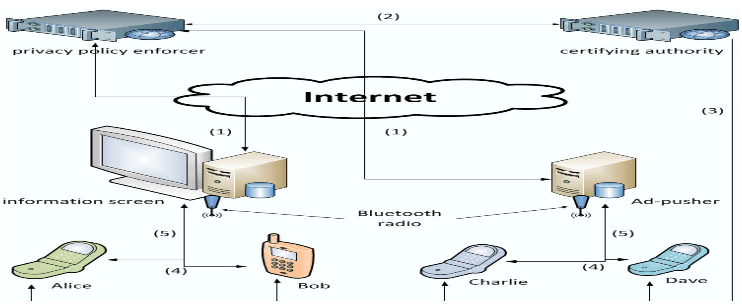


Fig. 1. Policy enforcement components

Initially SPs declare personal data they intends on soliciting from users in an unverified PP (uPP) and submit it to the EA*(1)*. uPP deemed to solicit and use data within EA's acceptable privacy practises (privacy threshold) are verified. EAs privacy practises are defined by authorities and set by enforcer administrator. EAs serves in similar capacity as a government/industry regulator to render significant punishment to misbehaving SPs (heavy fines, banned for periods of time). uPP exceeding EAs privacy threshold are returned to the SP with reason for failing and logged at the EA for privacy audit trail. Verified PP (vPP) incorporated with EA's credentials (signature, public keys) and forwarded to the CA *(2)* for signing and publishing. Certified vPP are fully functional PPs. EAs permits users and services to view PPs, SPs' audit trails, privacy thresholds and declined uPPs. Users download PP to their PTDs to facilitate interaction with specific SP *(3)*. SPs with verified and signed PP broadcasts their credentials and data request over Bluetooth *(4)*. PTDs with local copies of PPs perform credential validation and respond to solicitations transparently*(5)*.

3.1 Privacy Policy

The PP is an XML format file detailing a SP's credentials (service category, identity, name, description, certifying authority, owner, owners contacts and validity date) and data handling practises (data solicited, solicitation purpose, is the data mandatory/obligatory to service provision, third party disclosures, will this data be encrypted or not, duration of retention). EA administrator utilises privacy rules set privacy thresholds. Rules classifying data as either: personal identifiable, personal non-identifiable, contextual, preference, sensitive, and communicable attributes.

PPs bind the entire EAs system architecture together. The PP verification and PP certification binds the EA and CA. Publishing PPs and their installation by in PTDs, adds mobile terminals to the system. In this binding, only PTDs (users) maintain autonomy thereby reducing the association of a particular user with an instance of service use. Binding a group SPs with a single PP makes the architecture scalable. Suppose a shopping mall has 3 gates, placing Ad-pushers at each would require 3 different PPs. This cumbersome for users intending to use all Ad-pushers (download and install all 3 into their PTDs). Adopting group PP *(PPn)* as Figure 2 depicts, renders it possible to use all Ad-pushers by simply installing a group PP. Alternatively, a user can install a single PP *(PP1, PP2, PP3)* to avoid all SP instances. This scalability benefits PPs by easing large scale deployment, management and easing usability.

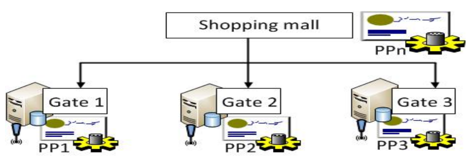


Fig. 2. Scalability in enforced privacy policies

3.2 Implementation

The SP is implemented using Python 2.6 and SQlite3 database on Ubuntu 10.04 Linux box. The SP and the remotely located EA interact over XMLRPC encrypted with elliptic curve cryptography. The EA administrator and users access the EA a over web interface. Composing the EA is Nginx web server, PHP, JavaScript and MySQL 5.1.41 database.

4 Conclusion

This study prototyped a mechanism of guaranteeing users privacy in ambient ubiquitous services. Privacy policies are implemented within an architecture consisting of an enforcement authority, a certifying authority and users' mobile terminals as personal trusted devices. The implementation addressed two key privacy challenges in ubiquitous services, personal data handling and accountability of adopted privacy measures. Practicality of the architecture are demonstrated as useable, scalable and easily manageable, thus a viable replacement for privacy seals in ambient ubiquitous environments. Our future works involves privacy challenges in a scenarios where users escalate to service providers roles for other services/users and their associated trust models.

References

1. Moores, T.T., Dhillon, G.: Do privacy seals in e-commerce really work? ACM Commun. 46(12), 265–271 (2003), http://doi.acm.org/10.1145/953460.953510
2. Moores, T.: Do consumers understand the role of privacy seals in e-commerce? ACM Commun., 86–91 (2005), http://doi.acm.org/10.1145/1047671.1047674
3. Kim, D.J., Steinfield, C., Lai, Y.-j.: Revisiting the role of web assurance seals in consumer trust. In: Proceedings of the 6th International Conference on Electronic Commerce, ICEC 2004, Delft, The Netherlands, pp. 280–287. ACM, New York (2004), http://doi.acm.org/10.1145/1052220.1052256
4. LaRose, R., Rifon, N.: Your privacy is assured-of being invaded. Web sites with and without privacy seals (March 2003), https://www.msu.edu/~larose/es2003post.htm (accessed February 4, 2010)
5. The World Wide Web Consortium, Preferences 1.0 (P3P1.0) Specification. W3C Recommendation. WC3 (2002), http://www.w3c.org/TR/
6. Mbanaso, U.M., Cooper, G.S., Chadwick, D.W., Proctor, S.: Privacy Preserving Trust Authorization Framework Using XACML. In: WOWMOM 2006: Proceedings of the 2006 International Symposium on World of Wireless, Mobile and Multimedia Networks, pp. 673–678. IEEE Computer Society, Washington (2006), http://dx.doi.org/10.1109/WOWMOM.2006.92
7. Saklikar, S., Saha, S.: Next steps for security assertion markup language (saml). In: Proceedings of the 2007 ACM Workshop on Secure Web Services, SWS 2007, Fairfax, Virginia, USA, pp. 52–65. ACM, New York (2007), http://doi.acm.org/10.1145/1314418.1314427
8. Serjantov, A.: A fresh look at the generalised mix framework. In: Borisov, N., Golle, P. (eds.) PET 2007. LNCS, vol. 4776, pp. 17–29. Springer, Heidelberg (2007)
9. Brown, P.W.: Digital signatures: can they be accepted as legal signatures in EDI? In: CCS 1993: Proceedings of the 1st ACM Conference on Computer and Communications Security, Fairfax, Virginia, United States, pp. 86–92. ACM, New York (1993), http://doi.acm.org/10.1145/168588.168598

A Concept for a First Communication Initiation for Ambient Intelligent Industrial Environments

Florian Floerchinger[1] and Marc Seissler[2]

[1] SmartFactoryKL, P.O. Box 3049,
D-67653 Kaiserslautern, Germany
floerchinger@smartfactory.de
[2] University of Kaiserslautern,
Gottlieb-Daimler Str. 42, D-67663 Kaiserslautern, Germany
Marc.Seissler@mv.uni-kl.de

Abstract. This paper presents a concept for an intuitive way of establishing physical communication connections between industrial field devices and an universal control device that supports several communication standards. It aims to stimulate the discussion about how a first communication initiation in ad-hoc linked environments has to look like. Key of the presented concept is a reference model that specifies all communication relevant information of a field device, attaching this information to the device, using an inexpensive passive data carrier. Furthermore, this reference model references data like multi-user behavior and user-rights, enabling a flexible user-centered device interaction which satisfies the needs of an industrial environment. With such a reference model most AMI technologies will be much more easy to use in ad-hoc linked environments. The concept presented in this paper is being tested and evaluated in the *SmartFactory*^{KL}, the intelligent factory of the future.

Keywords: Universal Interaction Device, SmartFactory^{KL}, context model, industrial communication, user-centered interaction, useML.

1 Introduction

Visionary research forecasts adaptive industrial environments linked up ad-hoc in the factory of tomorrow [1, 2]. Over the last years industry has recognized the demand for universal interaction devices which promise to enable the seamless control of different devices in a flexible industrial setting. In visionary future factories, devices may be connected without any configuration effort and the operator will be able to control all functions according to his qualification appropriately. Model-based development techniques will play an important role to enable the design of user-centered interaction devices that allow the adaption of the User Interface to the context of use, like the users current location [3]. Today, industrial reality is far away from this desired Plug & Play behavior. To let this vision become reality, a generic reference model is demanded which describes the different direct and indirect communication connections for industrial field devices. In addition, it has to be possible to reference information about operating-specific data, like user requirements, multi-user behavior and security demands to satisfy the needs of an industrial environment.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 270–274, 2010.

This reference model can be stored directly in form of a binary data representation on a passive data carrier – e.g. a Radio Frequency Identifier (RFID) – mounted directly on the field device. Reading this data carrier with the universal control device the field device will be identified and the communication will be established without any additional configuration effort after checking the operating-specific data (see Fig.1).

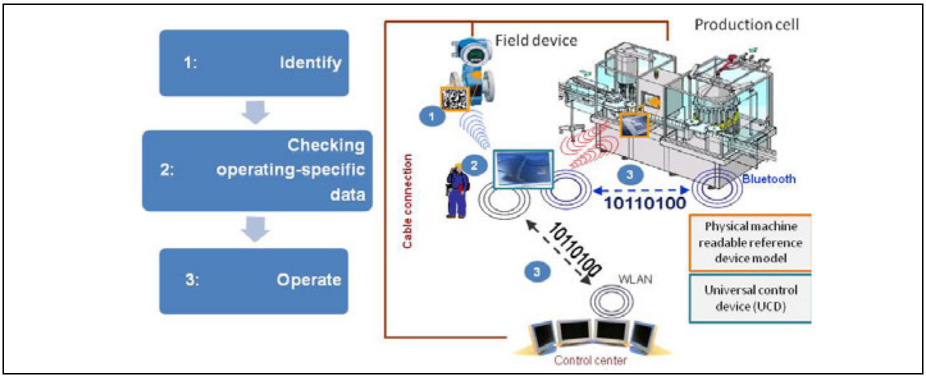


Fig. 1. Concept for an intuitive way of establishing physical communication connections

2 Preliminary Work and Lessons Learned – Universal Control Devices in Ambient Intelligent Production Environments

The number and complexity of technical devices in industrial plants are constantly growing. What is common are thousands of field devices and plant modules in a single facility. All these devices can be connected through different communication connections and have to be configured and managed by different workers. Without reducing the configuration effort these complex field devices will become less manageable which will have an impact on the flexibility of modern factories.

Examples for such a challenge are human machine interfaces (HMI) of industrial field devices. Since these devices often stem from multiple vendors, they possess user interfaces (UIs) of different types. Since the users are confronted with different UIs, operation errors often occur. Therefore workers can benefit from having an universal control device (UCD) which allows an uniform access of different devices.

To show the feasibility of an UCD within ambient intelligent production environments several demonstrators have been developed. In a first feasibility study several mobile phones have been used as an UCD to control a core module of the *SmartFactory*^{KL} demonstration plant via a proprietary Bluetooth connection [4].

Due to the very positive feedback of the initial UCD prototype and based on the lessons learned a new UCD – the unipo® UCP450 control device – has been developed in a follow-up project in collaboration with two industrial companies [5].

While the UCP450 allows a broad access of the various field devices located in the *SmartFactory*^{KL} it doesn't consider the context of use which is important to secure the safety respecting user-rights, multi-user behavior and real-time issues. Additionally this information allow the run-time adaption of the user interface. This challenge is currently addressed within the research project "GaBi" where a prototype of a runtime adaptive UCD – called SmartMote – for ambient intelligent production environments has been developed [6]. To enable the adaption of the user interface during runtime a model-based approach is used. Core of this approach is an use model, specified in the Useware Markup Language (useML), which allows the specification of detailed information about users' tasks, like task structure, task types, conditions and temporal relationships between the different tasks [7].

One lesson learned from all these projects is, that several preconditions have to be fulfilled to describe an adaptive UI in an ad-hoc combined industrial environment. First step is the unique identification of the field device which should be connected to. Subsequently a physical communication connection has to be established, which is rather complex with regard to all the different industrial communication standards. After establishing a communication connection questions dealing with user rights, multi-user behavior and security demands have to be answered before an adaptive UI can be built up and a worker is allowed to operate a device. Currently there is no known research project (as far as we know) which focuses on the question, how these context information sources can be referenced in real-time for the demands of industrial devices. Today, context information is widely used within static industrial environment that do not change over time. Future factories will need all information to be dynamically referenced in real-time from different sources (see Fig.2), because in an adaptive and flexible factory, production lines change quickly.

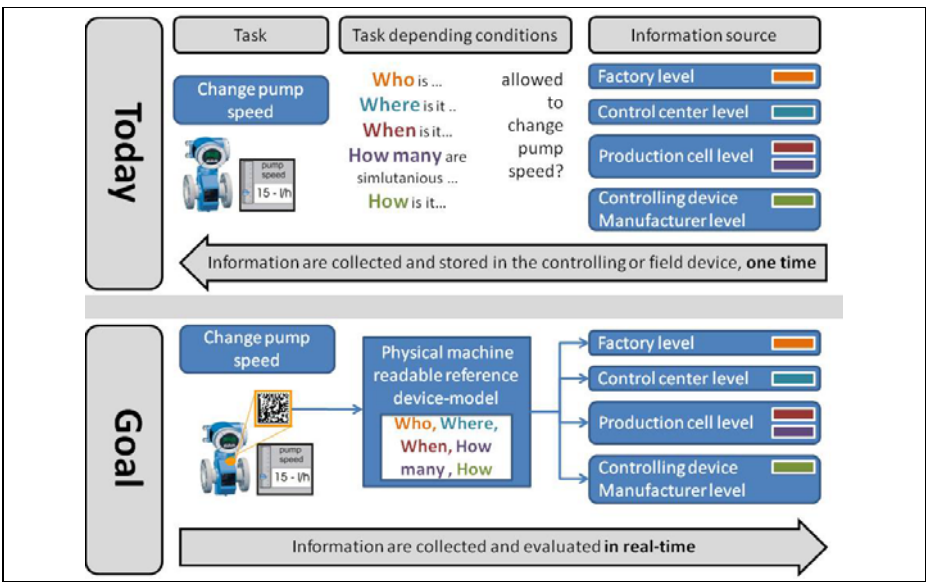


Fig. 2. Context information of the future generic device reference model

The generic reference model is one reasonable approach to store all relevant device and context references directly on the field device, which is extended by an inexpensive passive data carrier (e.g. re-writable like RFID, NFC or read only like Data matrix, QR Code).

3 Project Objectives and Goals

The goal of the project is to create a framework which allows a task-based initial communication to industrial field devices. The device operator should be supported the best way in his workflow, without considering how to configure the used devices. The **unique identification** of the field device is the first step to a generic device identification regarding the **needs of ad-hoc combined industrial environments**. The **description of the communication interface and the topology** are the next challenges, because a physical communication connection is a precondition for operating a field device. A generic model has to be developed which is able to reference the typical industrial interfaces e.g. ProfiNet, ProfiBus, RS232, RS 485, IWlan and Bluetooth. Establishing a communication must be enabled solely on the information stored in the generic model. Therefore common **communication-relevant attributes of widespread communication standards** have to be evaluated. On the basis of this information it will be possible to configure physical communication interfaces. To enable the automatic processing of the **generic reference model**, the identified attributes have to be stored in a machine-readable data format.

The generic reference model, which describes the device-specific communication attributes, is stored on a **passive data memory directly mounted on the field device**. In consideration of the research results regarding establishing secure communication channels using out of band-canals (OOB) [8, 9], passive data memories seem to be the most appropriate data carrier for industrial environments regarding the usability. Additionally the results of the "Semantic Product Memory"-project (SemProM) can be used as a first basis for this approach [10]. Using a SemProM the UCD can read the field device parameters, stored on the memory and establish a connection (see Fig.1). To show the feasibility of the concept a first prototypical implementation of a communication model interpreter has been integrated in the SmartMote universal control device. This prototype allows establishing a connection to different industrial field devices, using the communication information stored on RFID-tags.

4 Conclusions and Outlook

The generic device identification for an user-centered, task-based communication connection represents a further step towards a more intuitive human-machine-interaction in future factories. An initial communication reference model for describing the common attributes of communication interfaces has been developed and is actually evaluated in the *SmartFactory*^{KL}, with different industrial field devices from several vendors.

Future work will address the adaption of the UI based on the context information stored within a generic reference model. Therefore we focus on the useML-ID approach. The generic reference model will also support the adaptive useML-based UI. In this research approach the UI is built out of elementary use objects which all have an ID. Thinking of Fig.2 the information about e.g. multi-user behavior and user rights is spread all over the factory. This distributed information is linked to the useML-ID, i.e. that the conditions can be directly linked to the elementary use objects. Based on this information the use objects can be activated or deactivated in the generated visualization depending on the individual user rights in real-time. With the generic reference model it will be possible to connect unknown field devices in ad-hoc linked industrial environments and to calculate the accessibility of use objects depending directly to the context of use in real time.

References

1. Jovane, F., Westkamper, E., Williams, D.: The ManuFuture Road: Towards Competitive and Sustainable High-Adding-Value Manufacturing. Springer, Berlin (2009)
2. Zuehlke, D.: SmartFactory - A Vision becomes Reality. In: Keynote Papers of the 13th IFAC Symposium on Information Control Problems in Manufacturing (INCOM 2009), Moscow, Russian Federation, ICS / RAS, June 3-5 (2009)
3. Breiner, K., Maschino, O., Goerlich, D., Meixner, G., Zuehlke, D.: Run-Time Adaptation of a Universal User Interface for Ambient Intelligent Production Environments. In: Proc. of the 13th International Conference on Human-Computer Interaction (HCII) 2009, San Diego, USA. LNCS, vol. 5613, pp. 663–672. Springer, Heidelberg (2009)
4. Goerlich, D., Stephan, P., Quadflieg, J.: Demonstrating Remote Operation of Industrial Devices using Mobile Phones. In: Proc. of the 4th Intl. Conference on Mobile Technology, Applications and Systems, Singapore, pp. 482–485 (2007)
5. Meixner, G., Petersen, N., Koessling, H.: User Interaction Evolution in the SmartFactory^{KL}. In: Proc. of the 24th BCS International Conference on Human-Computer Interaction, Dundee, Scotland (2010)
6. Breiner, K., Maschino, O., Goerlich, D., Meixner, G.: Towards automatically interfacing application services integrated in an automated model-based user interface generation process. In: 4th Workshop on Model-Driven Development of Advanced User Interfaces (MDDAUI), Intelligent User Interfaces (IUI) 2009, Sanibel Island (2009)
7. Meixner, G., Seissler, M., Nahler, M.: Udit – A Graphical Editor For Task Models. In: Proc. of the 4th Int. Workshop on Model-Driven Development of Advanced User Interfaces (MDDAUI), Sanibel Island, USA, CEUR Workshop Proceedings, vol. 439 (2009)
8. McCune, J.M., Perrig, A., Reiter, M.K.: Seeing-is-believing: Using camera phones for human-verifiable authentication. In: IEEE Symposium on Security and Privacy (2005)
9. Kumar, A., Saxena, N., Tsudik, G., Uzun, E.: A comparative study of secure device pairing methods. In: Pervasive and Mobile Computing, vol. 5, pp. 734–749 (2009)
10. Wahlster, W., Kroener, A., Schneider, M., Baus, J.: Sharing Memories of Smart Products and Their Consumers in Instrumented Environments. IT – Information Technology, Oldenbourg. Special Issue on Ambient Intelligence 50(1), 45–50 (2008)

A Bluetooth-Based Device Management Platform for Smart Sensor Environment

Ivan Boon-Kiat Lim and Kin Choong Yow

School of Computer Engineering, Nanyang Technological University
Nanyang Avenue, Singapore 639798
ivan.bk.lim@pmail.ntu.edu.sg, kcyow@ntu.edu.sg

Abstract. In this paper, we propose the use of Bluetooth as the device management platform for the various embedded sensors and actuators in an ambient intelligent environment. We demonstrate the ease of adding Bluetooth capability to common sensor circuits (e.g. motion sensor circuit based on a pyroelectric infrared (PIR) sensor). A central logic application is proposed which controls the operation of controller devices, based on values returned by sensors via Bluetooth. The operation of devices depends on rules that are learnt from user behavior using an Elman recurrent neural network. Overall, Bluetooth has shown its potential in being used as a device management platform in an ambient intelligent environment, which allows sensors and controllers to be deployed even in locations where power sources are not readily available, by using battery power.

Keywords: Ambient Intelligence, Smart Sensor Environment, Bluetooth.

1 Introduction

In an ambient intelligent environment with numerous embedded sensing and actuating devices, a user should be able to set up his/her devices easily, without having to dismantle them for reconfiguration [1]. New sensors and controllers should also be easy to install without having to rewire or reconfigure the entire system. Due to this requirement, Bluetooth [2], with a low power consumption of 1mW, appears as a suitable wireless platform to implement an ambient intelligent environment.

In this paper, we describe the use of Bluetooth as a device management platform in an ambient intelligent environment, and demonstrate its feasibility in implementing an easy-to-setup array of sensors and controllers, which allows the expansion of sensors and controllers without messy wiring.

2 Existing Systems

Over the years, several standards and methods of implementing smart homes have evolved. The X-10 [3] is a low cost protocol for power line communications, with products and accessories produced by a large number of manufacturers. A key feature of the X-10 is that it does not need a central decision-making controller, and it uses

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 275–279, 2010.

the power line to transmit control signals to the devices using amplitude modulation. Another smart home technology standard is Echelon Corporation's LonWorks Networks [4], whose LonTalk protocol consists of a series of underlying protocols that allow intelligent communication among various devices on a network. However, the infrastructure requirements of such standards make it difficult for a user to add/remove sensors and actuators on an ad hoc basis.

3 A Bluetooth-Based Device Management Platform

An ambient intelligent environment needs to have some form of intelligent decision-making processor, in order to control and automate devices in the home. It needs to communicate with a variety of sensors to detect and guess what the user is doing, and also a variety of actuators to control the smart environment. We describe how we can add Bluetooth functionality to sensor and actuator circuitry, and to control them from a central logic controller application.

3.1 Sensor and Controller Hardware Design

The construction of Bluetooth-enabled sensors and controllers can be achieved by connecting the sensor and/or controller circuits to a microcontroller which has an I/O interface to a Bluetooth transceiver unit The analog signal generated by the sensor circuit will be converted to a digital signal and transmitted to the central computer application via Bluetooth. When control signals are received from the central logic application, the microcontroller will trigger the solid-state relay and turn a device on/off. Fig.1 shows the functional blocks of the sensor and controller hardware circuitry.

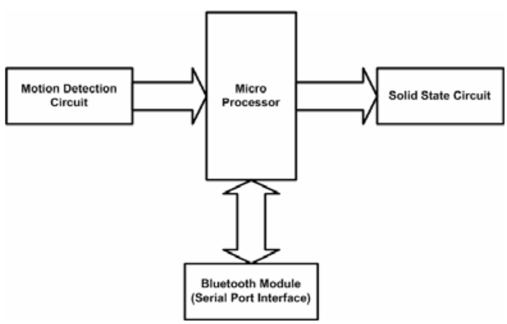
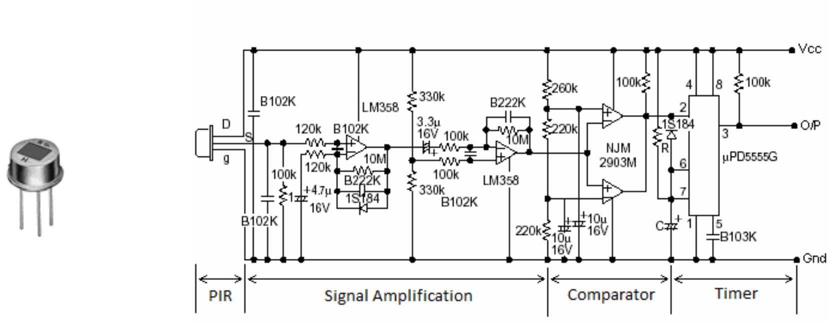


Fig. 1. Functional blocks of the sensor and controller hardware circuitry

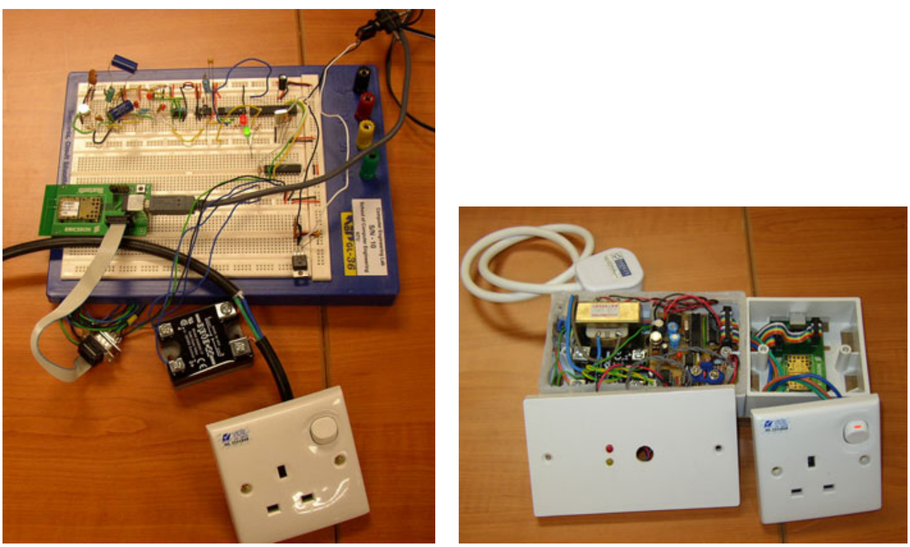
For example, to construct a motion sensor circuit to detect human presence in a room, we can make use of a pyroelectric infrared (PIR) sensor, coupled with a Fresnel lens to increase its coverage distance and angle. Fig. 2 shows the PIR sensor module and the motion sensor circuit which consists of three main stages. The first stage consists of operational amplifiers to amplify the waveform when movement is detected. The second stage consists of a window comparator, which converts the waveform into a pulse. This pulse is passed through the final stage, which consists of a 555-timer, which extends the pulse into a logic high output, whose width is determined by the value of the resistor and capacitor used.



(a) (b)

Fig. 2. (a) Pyroelectric infrared sensor module and (b) the sensor circuit design

The output of this motion sensor circuit is connected to the input port of a PIC 16F873 microcontroller, which sas a serial RS232C interface to an Ericsson Bluetooth Development Kit. The entire circuitry including the microcontroller can be assembled into a manageable box package, which is then fitted into a double power socket box. Fig. 3(b) shows the prototype with the contents exposed. When the cover is closed the entire package can be mounted on the wall.

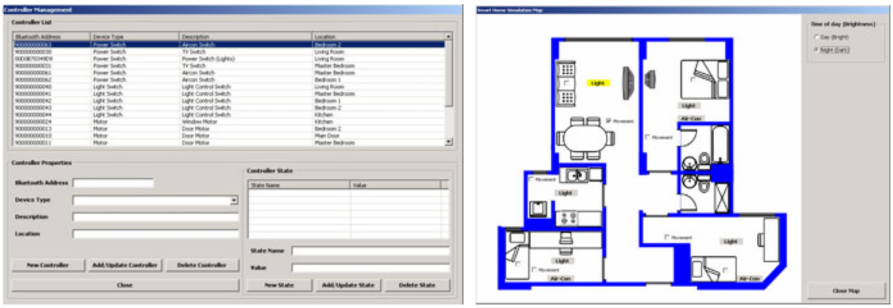


(a) (b)

Fig. 3. Hardware prototype (a) on a breadboard, (b) packaged in a socket box

3.2 Central Logic Controller Application Design

The central logic controller application is the “brain” of the entire intelligent environment and it allows a variety of sensors and devices to connect to it. Fig. 4 shows the screenshot of the sensor and controller management interface, and a map of the intelligent environment.



(a) (b)

Fig. 4. (a) Sensor and controller manager (b) Map of the intelligent environment

To overcome the Bluetooth limitation of only seven active slaves, the PC with the central logic controller application must assume the role of the master. When the application is started, it will initialize all sensors and controllers by first connecting to each of the slave devices, and setting them to parked mode. Subsequently, the master device can command a slave device to unpark itself to join the set of active slaves, or slave devices can request the master to unpark them so that they can transfer data to the master. Once the sensor device has finished transmitting its value, the master (PC) will put it in parked mode again, thus freeing up the communication channel.

When new sensors and controllers are added to the system, they can immediately be discovered by the central logic application and be used as part of the system. This ensures that the smart environment can continue to operate without having to restart the entire system and/or reinitialize all the devices.

3.3 Learning Logic Rules

The logic rules on how actuators will respond to a set of sensor inputs are learned using an Elman recurrent neural network [5]. The Elman RNN is a three-layer network, with the addition of a set of "context units" in the input layer. The context units always maintains a copy of the previous values of the hidden units, and thus the network can maintain a sort of state, allowing it to perform such tasks as sequence-prediction that are beyond the power of a standard multilayer perceptron.

During logic rule learning, the outputs of the sensors in the intelligent environment shown in Fig. 4(b) are fed directly to the inputs of the Elman RNN. Using supervised learning, the user can train the network to perform specific actions based on his position and his actions (e.g. if he moves from the bedroom to the toilet in the middle of the night, he may want his bedroom air-conditioner to remain running instead of turning itself off because no one is inside the bedroom).

4 Testing and Evaluation

In order to test the actual response and practical loading effect of the rules on the central logic application, a set of logic rules are learnt from the behavior of the user using the Elman recurrent neural network. With these rules implemented, the central logic application is tested for correctness, using the map to provide a visual indication of the controllers that are activated by the logic rules.

The results show that when the complete system is tested as a whole, the individual hardware and software units work according to the behaviour described in the previous sections. However, it is observed that when the central logic application sends data to the controller, the data sometimes do not get sent across due to microcontroller not being ready to receive data because it had just completed an earlier action. This problem can be adequately overcome by introducing a short delay to the central controller application before data is being sent out to the controller.

5 Conclusion

Ambient intelligent environment technology is gathering more and more interest among IT-savvy consumers, and Bluetooth is emerging as a technology for personal area networks. In this paper, we have merged these two concepts together to demonstrate the ease of adding Bluetooth capability to sensors and controller circuits, and the feasibility of the use of Bluetooth as a wireless technology in the implementation of a smart environment concept.

Bluetooth, being a low power consumption technology, also makes it suitable for deployment in battery-operated devices, which can then be installed in any part of the home, without the need to be close to a power source.

References

1. Aarts, E.H.L., de Ruyter, B.: New research perspectives on Ambient Intelligence. Journal of Ambient Intelligence and Smart Environments 1(1), 5–14 (2009)
2. Bluetooth Core Specifications, `http://www.bluetooth.com/English/Technology/Building/Pages/Specifcation.aspx` (retrieved March 17, 2010)
3. Wikipedia, X10 (industry standard), `http://en.wikipedia.org/wiki/X10_industry_standard` (retrieved February 20, 2010)
4. Echelon Corporation, LonWorks Technology Overview, `http://www.echelon.com/developers/lonworks/default.htm` (retrieved February 24 , 2010)
5. Pham, D.T., Liu, X.: Training of Elman networks and dynamic system modeling. International Journal of Systems Science 27(2), 221–226 (1996)

Investigation and Demonstration of Local Positioning System Using Ultrasonic Sensors for Wide Indoor Areas

Takashi Hada, Hikaru Sunaga, Masaki Akiyama,
Shigenori Ioroi, and Hiroshi Tanaka

Kanagawa Institute of Technology, Faculty of Information Technology,
1030 Shimo-ogino, Atsugi-shi, Kanagawa, Japan
{s1085046,s1085033,s075002}@cce.kanagawa-it.ac.jp,
{ioroi,h_tanaka}@ic.kanagawa-it.ac.jp

Abstract. This paper proposes an indoor positioning system for detecting the location of a moving object such as a person or a goods trolley in a wide indoor area. Conventional system requires synchronization between transmission and receiving unit. As a result, they include not only an ultrasonic part but also a radio part to maintain synchronization. We have developed a system that uses only ultrasonic waves, that is, an asynchronous system. The system configuration and sequence of operation are explained and the verification system which includes H8, PIC microprocessors and a PC for positioning calculation is described. It was confirmed that the proposed method is valid and the positioning error is within 100mm.

Keywords: Positioning System, Ultrasonic Waves, Asynchronous System, Positioning Error, Navigation.

1 Introduction

The GPS(Global Positioning System) is widely used for determining position in an outdoor area. It has become a universal system and the location information provided by GPS is widely used in navigation systems and many other service systems [1]. However, for indoor positioning applications, no common system has been established and various schemes have been investigated.

Positioning systems that are based on ultrasonic waves have been proposed. These systems area based on the trilateration, that is, the distance between two points which is measured by the difference in propagation time of radio waves and ultrasonic waves is used for positioning [2]. The propagation time of radio waves is neglected in this system. Alternatively, time synchronization between transmitting and receiving units can be implemented in order to detect the propagation time [3]. Consequently, these systems use radio devices, so they become somewhat complicated and must satisfy the radio regulations, which are different for different countries.

We use inverse GPS methods [4], that is, an arrangement of a transmitting and a receiving unit that is the reverse of that of GPS. This scheme does not require synchronization or information on the distance between a transmitting and a receiving unit. Therefore, this type of positioning system can be built using only ultrasonic units

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 280–284, 2010.

and this both simplifies the system configuration and avoids the need to comply with radio regulations. The verification model is designed to confirm the feasibility of such a positioning system which uses only ultrasonic units. The feasibility and positioning accuracy have been confirmed by the experiment.

2 Indoor Positioning Method

There are two basic techniques that can be used for a positioning system which uses ultrasonic waves. The relationship between the transmitting and receiving units is shown in Fig.1. The first technique is a synchronous system, in which the receiving unit knows the time of transmission of an ultrasonic signal. In this type of system, the distances between each receiving sensor and the transmitter are obtained from the propagation time and the velocity of sound. Some systems regard the time difference between the arrival of radio waves and ultrasonic waves sent from the transmitter at the same time, as the propagation time of the ultrasonic waves. Conventional ultrasonic positioning systems are based on this principle. In this scheme, the propagation times of t_0, t_1 and t_2 may be determined because synchronization between the transmitter and the receiving unit is maintained or the time difference between the received radio and ultrasonic waves is regarded as the propagation time of the ultrasonic waves.

The other technique is based on an asynchronous system, in which the clock of the transmission unit is independent of the receiving unit. Therefore, the propagation time cannot be obtained in this scheme. Only the delay times of each receiving sensors can be detected in this scheme. The delay time means the time that has elapsed since the first receiving sensor detected the ultrasonic wave. Consequently, three delay times, that is, t_1, t_2 and t_3 can be obtained from four receiving sensors as shown in Fig.1. The synchronization of all receiving sensors is maintained because they are included in one receiving unit, in which one clock governs the timing of the receiving unit. This scheme does not require synchronization between the transmitter and the receiving unit, and so it makes it possible to simplify the system configuration and to avoid the need to conform with radio regulations.

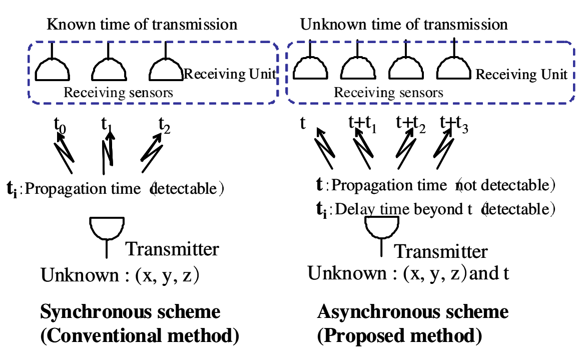


Fig. 1. Synchronous scheme and asynchronous scheme

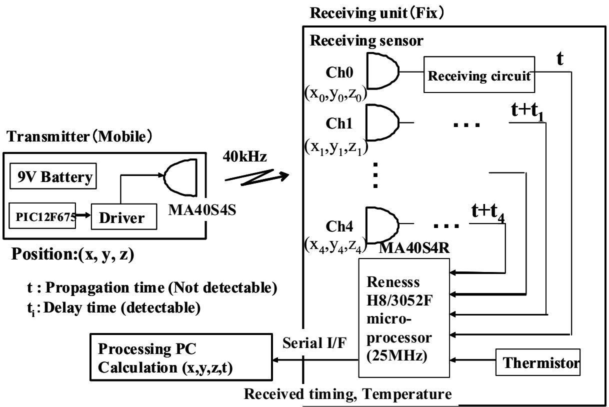


Fig. 2. Configuration of verification system

3 Positioning System Configuration and Its Sequence

The prototype system for verifying the scheme is shown in Fig.2. The positioning system consists of three units, that is, receiving unit, transmitter and signal processing PC. The Microchip PIC microprocessor is used to generate a 40 kHz on/off signal for the ultrasonic wave in the transmitter. The ultrasonic pulses are transmitted at a constant interval of 30ms and have a duration of 400μs.

The receiving unit is composed of receiving sensors, a Renesas H8/3052F microprocessor, the interface part for the PC, and the receiving circuits, each of which includes an amplifier and comparator. The number of receiving circuit is five, so five receiving sensors can be used for position calculation. The calculation is, in principle, possible with only four sensors but five sensors are used for certain position calculations, because there may be occasions where the relative positions of the transmitter and one receiving sensor make reception impossible. The output signals of the receiving sensors are amplified, and the comparators detect the arrival of the

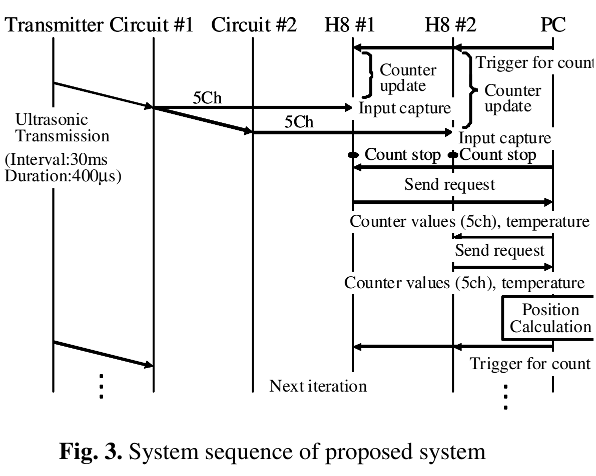
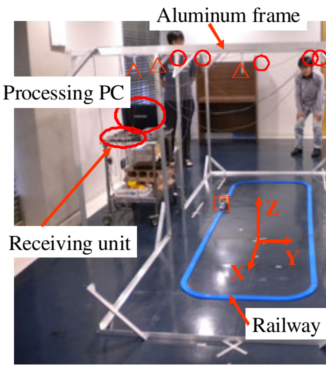
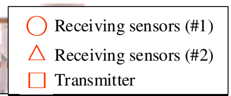


Fig. 3. System sequence of proposed system





		Position
Receiving unit #1	Sensor_0	(1000,0,1343)
	Sensor_1	(1500,500,1443)
	Sensor_2	(1500,500,1393)
	Sensor_3	(500,-500,1443)
	Sensor_4	(500,-500,1393)
Receiving unit #2	Sensor_0	(-1000,0,1343)
	Sensor_1	(-500,500,1443)
	Sensor_2	(-500,500,1393)
	Sensor_3	(-1500,-500,1443)
	Sensor_4	(-1500,-500,1393)

Fig. 4. Experiment configuration and sensor locations

ultrasonic waves at each sensor. The receiving sensors are Murata MA40S4R, whose diameter and height are 10mm and 7mm, respectively. The signal from the comparators in the circuits is captured by H8 processor ports and the counter values of H8 which indicate the receiving timing are sent to the PC via a serial interface. The thermistor is installed in order to monitor the ambient temperature.

The sequence of getting the received times, which are used to derive the delay time, from the receiving sensors via the H8 processors is indicated in Fig.3. The count start trigger to the two H8 processors is sent at the same time via the RS485 interface from the PC. The count values are held in each H8 micro processor when the associated sensor receives the ultrasonic wave. These values are sent to the PC and calculated to give the delay times. Five count values are obtained from the five ports that are connected to the five sensors.

The delay times t_i are calculated from these count values by taking into account the clock time of the H8 processors. Each count value is reduced by the count value of the first sensor to receive the ultrasonic waves, that is the minimum count value, in order to convert the counts to delay times. The position calculation is carried out taking account of the temperature, obtained as a voltage value from the thermistor, to determine the ultrasonic wave velocity c. The sequence from trigger to position calculation is repeated for subsequent iterations as shown in Fig.3. The transmitter and receiving units do not require synchronization, because only the delay time is used to calculate position.

4 Confirmation Experiment

The experimental configuration, the sensors installed in the frame and their locations are shown in Fig.4. The size of the railway track was used to evaluate the proposed scheme. Two identical receiving units were used for this experiment. The positioning results are shown in Fig.5. It was verified that positioning can be carried out without significant degradation of accuracy. The reason why the accuracy is degraded in this central region seems to be due to the drop the strength of the received ultrasonic wave. As the receiving signal becomes weaker, the fluctuation in the count values that are converted to time delay becomes large. However, the accuracy seems to be

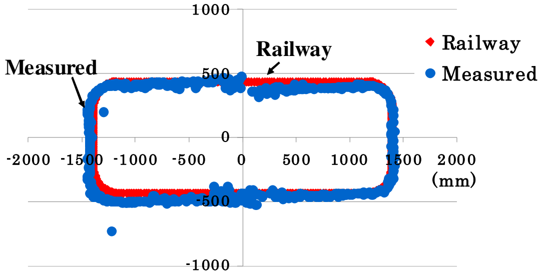


Fig. 5. Experiment result for proposed system

sufficient to detect the location of people who are walking. The peculiar result shown by the point at the lower left can be ignored by considering the moving velocity of the person. It has been verified that the proposed system architecture can be applied for any applications such as navigation in an indoor area, etc.

5 Conclusion

We have proposed an indoor positioning system which can be applied to the detection of the location of persons or moving vehicles in a factory. The main feature of the proposed system is that it uses only ultrasonic sensors. Therefore, it can be built more simply, compared with conventional systems and avoids the need to conform to radio regulations which differ from country to country. A verification system was built and the feasibility and accuracy of the proposed system have been confirmed by experiment. The accuracy is sufficient to monitor the location of person or goods vehicle in a factory, etc.

References

1. Khalid, M., Xiaodong, L., Yuji, H., Kinji, M.: Autonomous Pull-Push Community Construction Technology for High-Assurance. IEICE Transactions on Information and Systems E92.D(10), 1836–1846 (2009)
2. Priyantha, N., Miu, A., Balakrishnan, H., Teller, T.: The Cricket Compass for Context-aware Mobile Applications. In: Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom 2001), Rome, Italy, pp. 1–14 (July 2001)
3. Ward, A., Jones, A., Hopper, A.: A New Location Technique for the Active Office. IEEE Personal Communications 4(5), 42–47 (1997)
4. Homma, Y., Moriya, J., Hasegawa, T.: Comparative Experiments on Inverse GPS Based Systems Using Audible Sound or Ultrasonic Waves. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences J91-A(1), 139–142 (2008) (in Japanese)

Automatic Pedestrian Detection and Counting Applied to Urban Planning

Thomas Michelat[1], Nicolas Hueber[1], Pierre Raymond[1], Alexander Pichler[1], Pascal Schaal[2], and Bernard Dugaret[2]

[1] French-German Institute of Saint-Louis, ELSI,
5 rue du Général Cassagnou 68300 Saint-Louis, France
{Pierre.Raymond,Nicolas.Hueber,Thomas.Michelat}@isl.eu,
Alexander.Pichler@isl.eu
[2] Ville de Mulhouse, Pôle Voirie, rue Lefèvre 68100 Mulhouse, France
{Bernard.Dugaret,Pascal.Schaal}@mulhouse-alsace.fr

Abstract. The statistical knowledge of human flows in the streets is mandatory for urban planning. Today many cities use the expensive method of manual pedestrian counting, since there is no reliable automatic counting device. This project aims at achieving the first efficient, real-time, embedded and autonomous system that provides high-level data. Our first work focused on the development of a reliable counting method under Matlab™. On the basis of video sequences recorded in the city of Mulhouse we have validated the robustness of our approach.

Keywords: Pedestrian counting, Pedestrian detection, Background subtraction, Smart sensor, Global sensing, Embedded computing.

1 Introduction

Pedestrian flow statistics in the urban environment are essential to urban planning. They describe whether a novel layout up in the pedestrian area disturbs or facilitates the traffic. Therefore, the knowledge of human flow trends is relevant to optimize new urban planning, as well as the attractiveness and accessibility of city centers. Every town needs to frequently update these statistics. Today the standard and best method is based on manual counting. However, this method gives a limited survey of the flux statistics: only the direction and the number of people are evaluated, moreover, it requires a costly staff to collect these data.

Indoor automatic counting systems offer an efficient solution. Nevertheless, those systems suffer from a lack of portability in outdoor applications which require adaptation to large-area analysis and to environmental changes (luminosity, weather, camera jitter, etc.). In fact, their performance strongly decreases when they are used outdoors. We have checked the efficiency of a common counting device in a pedestrian area. This device is positioned by the side of the street at pedestrian level (1 m high) which is why it is unable to distinguish one person from a pedestrian row. The results gave strong unpredictable errors varying from 20% to 50%. Also its poor detection range limits the use of this device to narrow streets (< 4 m).

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 285–289, 2010.

This paper presents a robust solution to the outdoor constraints. On the basis of video analysis, we chose to embed the counting process directly after image acquisition and to output only high level data.

2 Methodology

2.1 Set-Up

A measurement campaign was conducted in a pedestrian street of the city of Mulhouse so as to determine the feasibility of automatically measuring the number of pedestrians and their direction by means of video analysis. The width of the street was about 15 m. In order to test our solution in real operational conditions, the camera was positioned on the side of a building along the street at various heights ranging between 8 and 12 m (Fig. 1). The pedestrian average flow was about 2200 people/hour (pph). This configuration induces a number of difficulties that must be overcome, e.g. some people are hidden by others, their height and shape change according to the perspective.

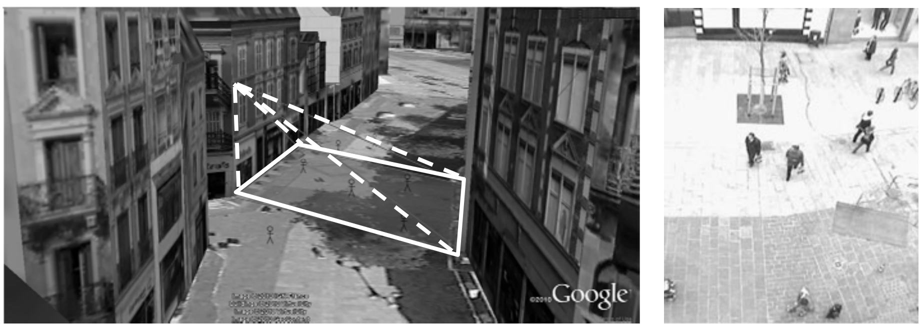


Fig. 1. Camera set-up and field of view (*left*); camera view (*right*)

2.2 Counting Method

The algorithm architecture shown in the figure 2 consists of two image processing methods: one for detecting and localising movement and the other for determining its direction. The movement is detected by evaluating local changes in the image along a line transversal to the pedestrian flow. It means that the background estimation is essential to an accurate detection. Many methods are given in the literature to model the background [1-3], but few offer a low computational cost which is our primary interest.

Based on the sigma-delta algorithm [4-5] and a local variance operator, the implemented background estimator is made robust to camera jitter, irrelevant motion such as swaying trees, change of illumination, and new static objects appearing on the scene. An example of background subtraction by this method is shown in figure 3. The background is estimated frame after frame.

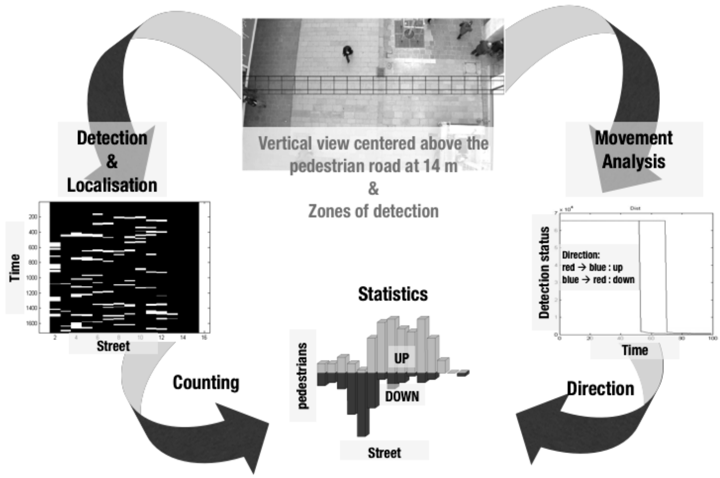


Fig. 2. Algorithm architecture

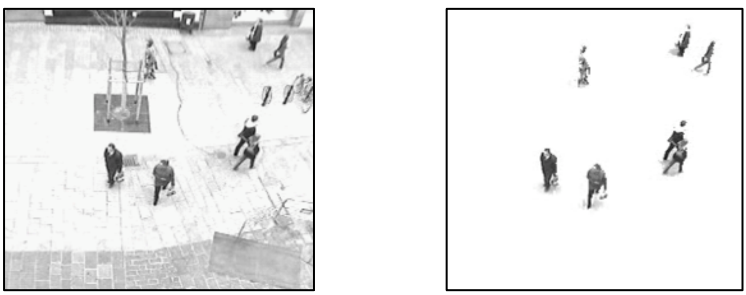


Fig. 3. Raw image (*left*) and the same image after background subtraction (*right*)

The direction is determined by processing the detection along a second line next to the first one. When a detection occurs at the same location on the two lines, the detection order between the lines indicates the direction of the pedestrian trajectory.

This algorithm does not involve high computational costs, and thus is well suited to real-time and autonomous devices; moreover, no calibration is needed, only the analysis area (pedestrian path) in the image has to be defined before starting the counting process.

3 Results

The recorded video sequences were carefully and manually analysed in order to build the exact counting database and to compare the data base results to our automated counting results. The confidence ratio for the counting and walking direction sensing is dependent on the duration of the analysed video sequence (Fig. 4). Starting from 92.2% at 2 minutes, it rapidly decreases, reaching 96.6% beyond 18 minutes in the case of a near-constant pedestrian flow around 2200 pph. An optimal height position of the camera was also measured at 10 m (± 1 m): the pedestrian image resolution was degraded beyond and perspective effects were stronger at a lower height, affecting the counting efficiency.

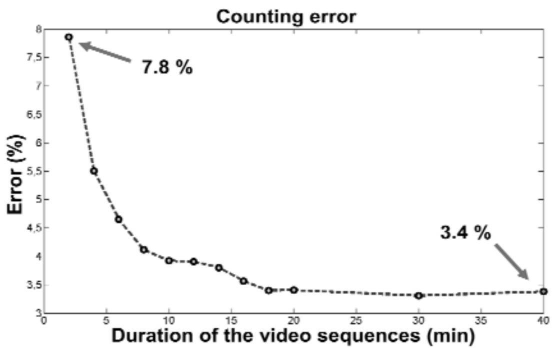


Fig. 4. Counting error versus video analysis time (pedestrian density E 2200 pph)

Despite the good confidence ratio, our algorithm remains slightly sensitive to perspective effects. In figure 5, the image is processed in two separate zones. The results show an underestimation for the upper part (zone 2) which is explained by the perspective effects, causing people to be hidden by others in pedestrian groups. The local change analysis along the line is extended to the whole image so as to map the pedestrian density (Fig. 6).

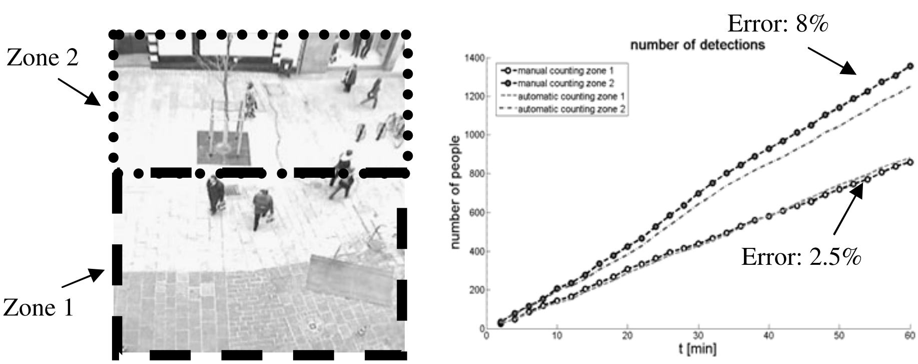


Fig. 5. Camera view and two analysis zones (*left*); counting curves (*right*) in the two zones for manual postprocessing counting and for automatic counting

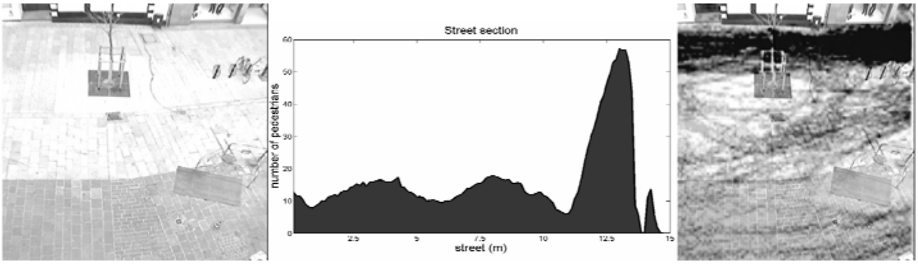


Fig. 6. Background estimation (*left*), pedestrian density profile along the street (*centre*) and image of the paths followed by pedestrians (*right, in black*)

4 Conclusion and Prospects

Our algorithm gave a very reliable and precise response in real conditions; moreover, it is less sensitive to perspective when pedestrians are hidden in a group. More thorough investigations will be carried out to evaluate the counting efficiency as a function of the crowd density [3]. Nevertheless, these results are sufficiently accurate to plan the algorithm implementation on a hardware prototype. This prototype will be designed to be autonomous, low-cost, real-time and to embed data processing in order to supply only the necessary statistics. At last, we will study the combination of this moving object detection with a recognition and classification process so as to determine whether the object is a pedestrian, a bike, a car, etc. Being able to map the density and trajectory of pedestrians opens the application field to many other topics of concern, e.g. crowd behavior analysis [6], attractiveness of shops, etc.

Acknowledgments. The authors would like to thank Yves Suma and the council workers for supporting the measurement campaigns. This project is gratefully funded by the city of Mulhouse and the French departmental council of Haut-Rhin.

References

1. Elgammal, A., Harwood, D., Davis, L.: Non-parametric Model for Background. In: Vernon, D. (ed.) ECCV 2000. LNCS, vol. 1843, pp. 751–767. Springer, Heidelberg (2000)
2. Benezeth, Y., Jodoin, P.M., Emile, B., Laurent, H., Rosenberger, C.: Review and Evaluation of Commonly-Implemented Background Subtraction Algorithms. IEEE, Los Alamitos (2008)
3. Mayo, Z., Tapamo, J.R.: Background Subtraction Survey for Highway Surveillance
4. Manzanera, A., Richefeu, J.: A new motion detection algorithm based on Σ-Δ background estimation. Pattern Recognition Letters 28, 320–328 (2007)
5. Mc Farlane, N.J.B., Schofield, C.P.: Segmentation and tracking of piglets in images. Machine Vision Applications 8, 187–193 (1995)
6. Moussaïd, M., Helbing, D., Garnier, S., Johansson, A., Combe, M., Theraulaz, G.: Experimental study of the behavioural mechanisms underlying self-organization in human crowds (2010)

This Is Me: Using Ambient Voice Patterns for In-Car Positioning

Michael Feld[1], Tim Schwartz[2], and Christian Müller[1]

[1] German Research Center for Artificial Intelligence (DFKI), Intelligent User Interfaces Department
[2] MMCI Cluster of Excellence, Saarland University
Saarbrücken, Germany

Abstract. With the range of services that can be accessed inside a car constantly increasing, so are the opportunities for personalizing the experience for both driver and other passengers. A main challenge however is to find out who is sitting where without asking explicitly. The solution presented in this paper combines two sources of information in a novel way: Ambient speech and mobile personal devices. The approach offers improved privacy by putting the user in control, and it does not require specialized positioning technologies such as RFID. In a data-driven evaluation, we confirm that the accuracy is sufficient to support a ten-speaker scenario in practice.

Keywords: Positioning, Speaker Recognition, Automotive HMI, User Modeling, Personalization.

1 Introduction

[6] introduces a scheme of positioning systems, which distinguishes exocentric and egocentric approaches. With exocentric approaches, a device carried by the user sends a unique signal to the environment that is used to calculate the current position of the device. The result is then sent back to the user like it is the case with cell-phone positioning based on cell-id. GPS is an example of the other category, egocentric: the device receives data from the environment and calculates the position by itself. This has clear advantages with respect to privacy, since the users can control whether they want to share their position information with a third-party, as opposed to RFID based solutions commonly used in asset tracking [4]. The system described in this paper is a hybrid approach, since sensors – in this case microphones – are installed in the car and pick up signals generated by the user, e.g. casual conversations or voice commands. With this information the car's system can determine which seats are taken, but it cannot identify *who* is sitting there. The information that enables this identification is stored on the user's personal device. Thus, the device is the link between the location information (provided by the car) and the user information (stored on the device) (see Fig. 2 (*left*)). By combining an ambient information source with

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 290–294, 2010.

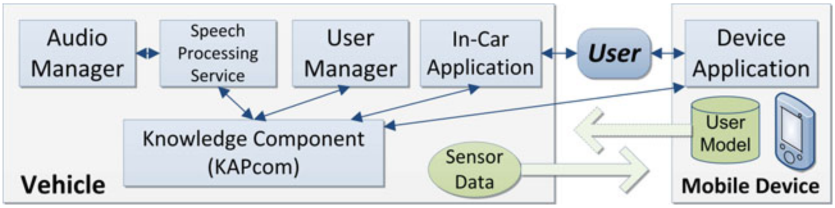


Fig. 1. The In-Car Positioning Architecture. Components are running either in the *vehicle* or *mobile device* context.

existing technology, positioning can occur in a fully automated and non-intrusive way.

Knowing what seat a passenger occupies enables a multitude of possibilities for personalization. One safety-related application are adaptive driver warnings [1]. Other features such as controlling car comfort functions benefit as well from knowing who sits where (e.g. temperature being configured to the user's preferred setting).

In the remainder of this paper, the in-car positioning approach based on ambient speech is described from a technical point of view. Since the reliability of such a system strongly depends on the accuracy of the speaker recognition process, empirical data is provided in Section 3. The major question is: how much speech do we need to reliably make that decision? As we will show, the system needs approximately one second of speech.

2 The Proposed In-Car Positioning Approach

The actual architecture of the proposed system consists of several hardware and software components that have all been installed into a real demonstration vehicle (Mercedes R-Class). The hardware set-up consists of two 7" screens (front seats located, center console) as well as two 10" screens (rear seats, head rest). All seats are equipped with hi-fidelity directional microphones (*Sennheiser ME 105-NI*). A Wireless LAN allows nomadic devices to easily join the car network. In our case, several Windows Mobile-based smartphones and PDAs represent the drivers' personal devices running the mobile client software.

Fig. 1 illustrates the software architecture. The central component that links together the other modules is the *Knowledge, Adaptation and Personalization component (KAPcom)*. It is a multi-purpose service maintaining a knowledge base with an automotive-domain ontology. Different applications can retrieve and store knowledge, e.g. user preferences or traffic information. In the positioning scenario, *KAPcom* is used as a "blackboard" – a design pattern for problem solving often applied in Artificial Intelligence [2]). On the data acquisition side, an *Audio Manager* obtains raw audio data from the microphones. It then streams the data to a *Speech Processing Service* responsible for detecting speech and – if speech is present – for computing a set of characteristic high-level audio

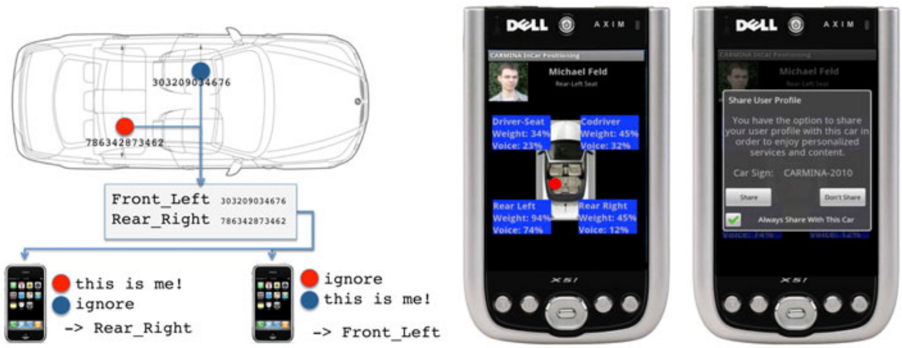


Fig. 2. *Left:* In-Car positioning. Sensor data in the environment (car) is transmitted to the user device(s). There, it is used as a key in order to find out what value belongs to the respective user. *Middle*: The system has identified its owner as the passenger on the rear left seat. *Right*: After positioning, the system asks whether the user wants to share their data with the car.

features, the so-called voiceprint (see also Section 3). This evidence is stored in the knowledge base together with the information on which seat it was recorded.

When a mobile phone or PDA connects to the car's network, it automatically starts receiving notifications using a mobile client application when new data become available. Given the user profile that is stored on the personal device, it can then locally perform the detection step with the sensor data from any seat, i.e. determine in how far a given voiceprint matches the signature stored in the user profile. Data is aggregated until a match can be confirmed with reasonable confidence. Afterwards, the device owner's location inside the car is known and the client can use this knowledge to perform device-side adaptation, such as changing the phone's ringing scheme and other notifications when the user is the driver. The current UI on the device (Fig. 2 (*middle*)) provides a visualization of the match likelihood for the evidence coming from each seat. The device owner can additionally decide to share his identity with the car, thereby allowing applications running on its system to also take advantage of personalization. Another component running in the car: the User Manager manages the association of user profiles to seats based on evidence and authentication data.

The proposed approach differs from other positioning methods in that it does not require the user's device to broadcast information, hence adopting the privacy advantage of egocentric positioning. However, when the positioning data is used to adapt car functions, information sharing concepts are needed to deal with the potential issues. Most critically, users have the option to block all data sharing in the first place on their device. When information should be shared, the situation may depend on whether the user is the owner of or a regular passenger in the car, or not. If so, the user profile can either be automatically transmitted by default without user interaction, or it can be stored on the car or even an off-board service and be only activated by a token on a device.

3 Generating and Testing Speaker Models

This work employs a speaker detection method that is motivated by creating a resource-efficient detection scheme for the mobile device implementation and an interchangeable model / voiceprint representation.

A basic GMM-UBM approach [5] makes the core of the recognizer. In this approach, speaker models are represented by Gaussian Mixture Models (GMMs), each model consisting of a fixed number of Gaussian probability distributions (in our case 1024) for each feature. The features are 12 coefficients of the mel-frequency cepstrum (MFCCs) indicating the short-term power spectrum of the voice. The GMMs are trained on pre-recorded speech and stored on the device in a compact binary format of approx. 200 KB each. In addition, a "background model" is trained from the speech of all speakers. The tool used to facilitate the training process is the speaker classification framework [3].

For any given audio stream from a live system and user model, a likelihood score can be computed that indicates to what extent the speech, which is segmented into frames, matches the model. The final acceptance of a speaker model occurs when a predefined minimum amount of material matches the speaker model more than the background model. In order to reduce sensitivity to noise and silence, a filter for voiced frames has been applied to all audio.

For training and evaluation, a corpus was recorded inside the standing vehicle described earlier. The corpus consists of ten adult speakers (7 male, 3 female) and a total of 76 minutes of material categorized by different conditions (seat, read vs. spontaneous, overlapping, doors closed vs. open). The training set consists of 30 minutes of speech, while the evaluation corpus uses 20 seconds (2000 frames) per speaker. There are two types of performance measures applied in this task: accuracy and error rates. The *accuracy* describes what percentage of test samples are assigned to the correct speaker, i.e. it measures how well the system can *identify* the correct speaker out of the full set of (in this case) ten speakers. The second category quantifies the errors a system can make when it tries to *detect* whether a test sample is from a given target speaker, which are *misses* and *false alarms*. A mobile device as in our scenario only has a single user profile stored and hence performs the *detection* task. The false alarm rate is possibly the most critical measure here because it tells in how many cases the device will incorrectly report a match. Our evaluation therefore takes into account different numbers of frames over which scores were averaged.

Fig. 3 illustrates the results. In a single-frame scenario (10 ms of speech), already 62.7% of frames are classified correctly at a chance level of 10%. However, it is clearly evident that recognition rates improve significantly when more speech becomes available. At 64 frames, which corresponds to 640ms of pitched voice (roughly a second of ordinary speech), almost 100% accuracy are achieved in this test. The corresponding chance of a false alarm is still 2.8%. It can be lowered to 1.6% when using approx. one fourth of the eval material for averaging.

Frames	Duration	Accuracy	P_{err}	P_{miss}	P_{fa}
1	0.01 s	62.74%	15.86%	13.65%	18.06%
2	0.02 s	73.53%	11.60%	9.34%	13.85%
4	0.04 s	84.39%	7.83%	5.45%	10.21%
8	0.08 s	92.76%	5.02%	2.89%	7.16%
16	0.16 s	97.19%	3.30%	1.46%	5.15%
32	0.32 s	99.19%	2.19%	0.62%	3.77%
64	0.64 s	99.91%	1.50%	0.18%	2.81%
128	1.28 s	100.00%	1.10%	0.00%	2.20%
256	2.56 s	100.00%	0.94%	0.00%	1.88%
512	5.12 s	100.00%	0.78%	0.00%	1.56%

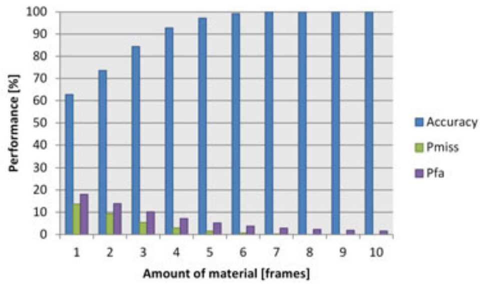


Fig. 3. Results of the in-car speaker recognition on a test corpus. The error rate P_{err} is defined as the mean of P_{miss} and P_{fa}.

4 Conclusion and Future Work

As documented in the previous section, the approach presented herein can indeed be used to rather quickly and reliably determine speakers' positions in a setting with a manageable number of speakers. Yet, since positioning/identification solely on voiceprints has some drawbacks, e.g. the passengers have to speak in order to be positioned, we will expand the presented system with weight sensors. Here, the basic idea is, that the users' profiles will be enriched by their body weight data. This value will of course change, even throughout the day, so appropriate estimation models will have to be found and evaluated. Both approaches, speaker recognition and weight measurements, will then be combined according to the "Always Best Positioned" paradigm, which was proposed in [6].

References

1. Brouwer, R.F.T., Hoedemaeker, M., Neerincx, M.A.: Adaptive interfaces in driving. In: HCI, vol. (16), pp. 13–19 (2009)
2. Corkill, D.D.: Blackboard systems. AI Expert 6, 40–47 (1991)
3. Feld, M., Müller, C.: An Integrated Development Environment for Speech-Based Classification. In: Proceedings of the 13th International Conference "Speech and Computer" SPECOM 2009, St. Petersburg, Russia, pp. 443–447 (June 2009)
4. Hayashi, H., Tsubaki, T., Ogawa, T., Shimizu, M.: Asset tracking system using long-life active rfid tags. NTT Technical Review 1(9), 19–26 (2003)
5. Reynolds, D.A., Quatieri, T.F., Dunn, R.B.: Speaker verification using adapted gaussian mixture models. In: Digital Signal Processing, p. 2000 (2000)
6. Schwartz, T., Stahl, C., Baus, J., Wahlster, W.: Seamless Resource-Adaptive Navigation. In: Resource-Adaptive Cognitive Processes, pp. 239–265. Cognitive Technologies, Springer (2010)

Selective Delivery of Points of Interest

Nuno Gil Fonseca[1,2], Luís Rente[1], and Carlos Bento[1]

[1] Centro de Informática e Sistemas da Universidade de Coimbra, Portugal
[2] Escola Superior de Tecnologia e Gestão de Oliveira do Hospital, Portugal

Abstract. In our daily life, we are increasingly surrounded by devices that expose us to quantities of information well behind our cognitive capabilities. To overcome this problem various authors propose mechanisms capable of selecting only the most relevant pieces of information.

In this paper, we propose an approach for information selection based on the concepts of relevance, selective attention and diversity. The idea is to select the most promising items in terms of surprise and usefulness and dismiss those that are less promising.

We illustrate our approach with an example of an application for restaurant selection and show the first results from an initial evaluation of this system.

1 Introduction

Evolution already provided humans with selective attention components that indicate which few aspects of the world are significant to a particular problem at hand. However, the overwhelming amounts of irrelevant information that they receive and need to process may compromise the person's performance.

This phenomenon, sometimes referred as information overload [1], is specially problematic (and eventually dangerous) if the person is performing attention demanding tasks, like driving.

Given the wealth of information in conjunction with humans real-time multi-task constraints, devices that incorporate artificial selective attention mechanisms with the aim of selecting only the relevant information are fundamental.

One example of applications where the information overload occurs are the POI-delivery applications, like the ones found on navigation devices and smartphones. These applications provide information about the nearest Points of Interest (POI), namely, gas stations, restaurants, hospitals, pharmacies, etc.

Most of the existing applications provide POIs regardless the user's real intention and of what s/he already knows, leading to the presentation of large amounts of irrelevant items that difficult the process of finding the desired one.

In the next sections, we will present some of the current approaches to overcome this problem, detailed information on our approach and some preliminary results. Finally, we present some conclusions and ideas for future improvements.

2 Current Approaches

To overcome the problem of having large amounts of location-based items to display, some mechanisms were already developed.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 295–299, 2010.

In most cases, the POIs are simply sorted according to specific criteria like popularity or spatial distance and then first n are displayed to the user.

Several approaches determine which POIs should be displayed to the user based on the relevance of each POI for the current context [2][3][5].

If the items are to be displayed on top of a map, several authors suggest the use of different generalization operations (selection, simplication, aggregation, typification and displacement) to make the maps more readable [6][4][5].

Although the presented solutions may be effective on selecting the most relevant items, they miss an important issue, which is the familiarity that the user already has with the POIs. In some cases, the user may be considerably familiar with most of the presented POIs, so, presenting them will eventually be useless.

3 Our Approach

Our approach has the goal of autonomously select the most relevant and unfamiliar items, preventing the delivery of irrelevant or unwanted information. This is achieved by employing information on the users context and intention in three consecutive steps: relevance, surprise and diversity.

The user's context is defined by his/her location and time, while the user intention/goal is input by the user as a set of attributes.

The relevance of each POI for the current context/intention is defined by the weighted sum of the distances between the k target attributes input by the user (T_j) and the ones associated to each POI (P_j).

$$Relevance(POI) = \sum((1 - Distance(T_j, P_j)) \times w_k), w_k \in]0, 1[\tag{1}$$

The POIs with relevance values greater than a specific threshold α are selected for the next step.

Inspired on natural selective attention [7], we propose the integration of an artificial selective attention (ASA) component relying on the cognitive model of surprise proposed by Macedo and Cardoso [8] to select the POIs that are most unfamiliar to the user.

The idea, is to determine the surprise felt by the user when confronted with a specific POI within a specific context and then select the items with the highest surprise values, which are the ones that the user most probably is unaware.

The surprise of each POI, $S(X)$, is calculated based on the probability that it has of being the desired one. This probability is defined by the user's appropriation level [9] which defines the relation that the user had with the POI on previous usages of the system.

$$S(X) = log_2(1 + P(Y) - P(X)) \tag{2}$$

Where X represents each individual POI and Y the POI with the highest probability of being the desired one.

The POIs with surprise values higher than the threshold μ (the ones with which the user had less contact in the past, or at least the less significant contacts) are selected to the next step.

For now, we are not having into account the fact that the low appropriation level for a given POI may result on the users dislike for that POI. For these types of situations, an exclusion mechanism should be included.

The final step of the process consists on selecting the most diverse POIs, to avoid selecting very similar POIs. The approach used was inspired by Gago et al. [10] and consists on determining the difference between each POI and all the others based on their attributes, and consecutively select the most diverse one.

The process stops after n POIs are selected or when there are no more candidate POIs with an average distance to the selected ones above a β threshold.

Finally, the most diverse POIs are displayed to the user.

To best meet the user's preferences, the system has some parameters that can be adjusted (the thresholds: α, μ and β or n).

4 Preliminary Results

To evaluate the proposed approach, we have conducted a simple experiment that consisted on using a smartphone with two different POI-delivery applications to obtain information about the nearest restaurants. The first application display all the retrieved POIs on a map without any kind of filtering (app. A) and the other will filter the POIs using our approach (app. B). In Fig. 1 are depicted the two applications used in the evaluation.

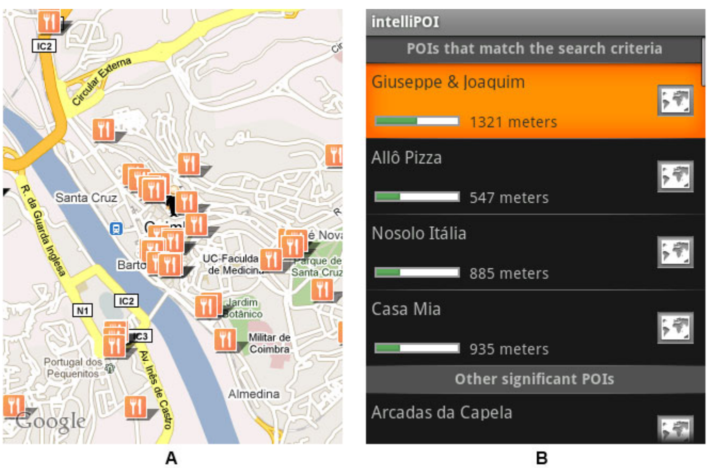


Fig. 1. The two applications used in the evaluation

The test was performed by 22 participants (17 male and 5 female) with an average age of 25,4. All of them have previously used POI-delivery applications.

We were specially interested in measuring two particular aspects: the ease of finding the desired restaurant and the time spent on that task. The participants rated the easiness in one 0-10 scale (being 10 the best result) and the time was

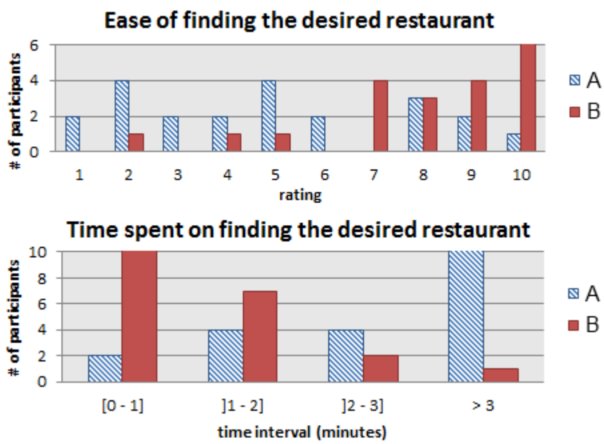


Fig. 2. Results on the ease of finding the desired item and the time spent on that task

measured using the application itself. On Fig. 2 we may observe the obtained results.

From these results we can see that most users rated the application B with 7 to 10 in opposition to the ratings of the application A, meaning that B was easier to use. It's also clear, that when using B the users were able to find the desired restaurant faster than when using A.

Unfortunately, due to time and logistic constrains, we weren't able to fully evaluate the effect of surprise on the delivered POIs. The participants used the applications only a couple of times and, due to it's nature, the effect of surprise would be noticed after several usages. Nevertheless, the tests provided us with a good idea on how the users react to the new approach in terms of the graphical interface. Most of them mentioned that the interface contributed to increase their performance.

After using both applications, all participants said that they would like to use the application B on their daily life.

5 Conclusions and Future Work

There are no doubts that POI delivery applications can help humans perform better in their daily tasks. However, it's also true that while using these applications, the users may spend large amounts of time filtering irrelevant and already known information.

With this in mind, we propose a new approach that uses information about the user's previous interactions with the system to provide relevant and unfamiliar items.

The presented results seem to be consistent with what we expected: eliminating irrelevant information may increase the user's performance. However, we are aware that the approach has still to be validated by a larger and more comprehensive group of users. It is also imperative, that the tests may last a larger period of time, so that the effect of surprise can in fact be noticed.

Although in this case, we have decided to deploy the proposed selective attention mechanism in the context of POI-delivery, it can be also applied on various other contexts involving delivery of information.

Besides performing new tests and improving the relevance, surprise and diversity methods/algorithms, one aspect that will deserve our attention in the near future concerns automatically determine the user's intention to proactively select and provide him/her relevant information for the particular context/intention.

References

1. O'Connell, N.: Interruption overload. Strategic Direction 24(10), 3–5 (2008)
2. Keim, D.A., Kriegel, H.P.: VisDB: Database Exploration Using Multidimensional Visualization. IEEE Computer Graphics and Applications 14(5), 40–49 (1994)
3. Sabbata, S., Reichenbacher, T.: A Probabilistic Model of Geographic Relevance. In: GIR 2010, Zurich, Switzerland (2010)
4. Burigat, S., Chitttaro, L.: Decluttering of Icons based on Aggregation in Mobile Maps. In: Map-based Mobile Services - Design, Interaction and Usability, pp. 13–32. Springer, Heidelberg (2008)
5. Pombinho, P., Carmo, M.B., Afonso, A.P.: Evaluation of Overcluttering Prevention Techniques for Mobile Devices. In: IV 2009 (2009)
6. Edwardes, A., Burghardt, D., Weibe, R.: Portrayal and Generalisation of Point Maps for Mobile Information Services. In: Map-based Mobile Services, pp. 11–30. Springer, Berlin (2005)
7. Driver, J.: A selective review of selective attention research from the past century. Br. J. Psychol. 92, 53–78 (2001)
8. Macedo, L., Reisenzein, R., Cardoso, A.: Modeling Forms of Surprise in Artificial Agents: Empirical and Theoretical Study of Surprise Functions. In: Proceedings of the 26th Annual Conference of the Cognitive Science Society (2004)
9. Fonseca, N.G., Rente, L., Bento, C.: Artificial Selective Attention for In-Vehicle Information Systems. In: Proceedings of the 2nd European Conference on Human Centred Design for Intelligent Transport Systems, Berlin (2010)
10. Gago, P., Bento, C.: A Metric for Selection of the Most Promising Rules. In: Zytkow, J., Quafafou, M. (eds.) PKDD 1998. LNCS (LNAI), vol. 1510, pp. 19–27. Springer, Heidelberg (1998)

Ambient Intelligence Research Landscapes: Introduction and Overview

Norbert Streitz

Smart Future Initiative
Konrad-Zuse-Str. 43, 60438 Frankfurt, Germany
norbert.streitz@smart-future.net
www.smart-future.net

Abstract. This paper starts out by introducing the "Landscapes" category at the Joint International Conference on Ambient Intelligence (AmI-2010) and provides an overview over the two sessions. The main part of the paper presents a framework for the role of Ambient Intelligence in the development of the cities of the future. This includes the integration of real and virtual worlds resulting in Hybrid Cities and their transformation into Smart Cities. In the context, it is argued that the technological development has to be monitored by guidelines and goals for maintaining and improving the quality of life leading to what is called Humane Cities, addressing, e.g., social awareness and privacy, trust and identity. The paper closes with proposals for a future research agenda.

Keywords: landscapes, ambient intelligence, smart environments, smart city, hybrid city, ubiquitous city, humane city, research agenda.

1 Introduction

"Landscapes" is a new category for presentations that was introduced for the first time at the European Conference on Ambient Intelligence (AmI-2009) in Salzburg, Austria. Landscapes do not fall in the classical category of research papers or posters but provide a comprehensive perspective. The motivation and purpose is to provide an international forum for presenting ongoing and new initiatives and activities, overviews over significant projects, test beds, laboratories and innovative organizations in the field of Ambient Intelligence. Landscapes are meant to help building a mental model of the field as well as identifying trends and new directions in research as well as in development.

Due to the success at AmI-2009, Landscapes are also part of AmI-2010 in Malaga. There are two Landscapes sessions with six presentations in total. The first session starts with an overview by the Landscapes Chair Norbert Streitz (Smart Future Initiative, Germany), introducing selected strategic issues for Ambient Intelligence, its role for the development of hybrid smart cities in the framework of creating a Humane City of the Future, and a set of research lines forming a new research agenda in the area of Ambient Intelligence Environments. This introduction and overview is followed by two papers sharing a common theme: Ambient Assisted Living (AAL),

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 300–303, 2010.

an area that receives increasingly attention. Reiner Wichert from the Fraunhofer-Institute IGD in Darmstadt, Germany, discusses the “Challenges and Limitations of Intelligent Ambient Assisted Living Environments”, presenting also the Fraunhofer Alliance Ambient Assisted Living which has the goal of building a universal modular platform for AAL and Personal Health applications. This is followed by Jochen Frey and his colleagues presenting “The DFKI Competence Center for Ambient Assisted Living” in Saarbrücken, Germany. It is a cross-project and cross-department virtual organization addressing the range from multimodal speech dialog systems to fully instrumented environments.

The second session has three papers dealing with ambient intelligence in retail and shopping, and the context of cars and factory. Carsten Magerkurth and his colleagues from SAP Research in St. Gallen, Switzerland, presents “Intersecting the Architecture of the Internet of Things with the Future Retail Industry”. Their approach combines an environment of co-located academic education (Campus-Based Engineering Centers) with the concept of living laboratories in which real-world prototypes and systems are rigorously tested in the Future Retail Center (FRC). This is followed by Evert van Loenen and his colleagues from Philips Research in Eindhoven, The Netherlands, discussing “The Role of ExperienceLab in Professional Domain Ambient Intelligence Research”. This paper goes beyond the consumer domain and presents professional domains such as retail, hospitality and healthcare, where different stakeholders play an important role. The session is concluded by Thomas Grill and his colleagues presenting “The Christian Doppler Laboratory on Contextual Interfaces” at the University of Salzburg in Austria. They discuss the implications of different notions of “context” and illustrate them in three domains: ambient persuasion, factory, and cars.

2 Towards Smart Ecosystems

The evolution towards a future information and knowledge society is characterized by the development of personalized individual as well as collective services that exploit new qualities of infrastructures and components situated in smart environments. They are based on a range of ubiquitous and pervasive communication networks providing ambient computing and communication at multiple levels. The collective services are provided by a very large number of "invisible" small computing components embedded into our environment. They will interact with and being used by multiple users in a wide range of dynamically changing situations. In addition, this heterogeneous collection of devices will be supported by an “infrastructure” of intelligent sensors and actuators embedded in our homes, offices, hospitals, public spaces, leisure environments providing the raw data (and active responses) needed for a wide range of smart services. Furthermore, new and innovative interaction techniques are being provided that integrate tangible and mixed reality interaction. In this way, the usage and interaction experience of users will be more holistic and intuitive than today. It is anticipated that economics will drive this technology to evolve from a large variety of specialized components to a small number of universal, extremely small and low-cost components that can be embedded in a variety of materials. Thus, we will be provided with a computing, communication, sensing and interaction “substrate” for systems and services. We can characterize them as "smart ecosystems" in order to emphasize the seamless integration of the components, their smooth interaction, the "equilibrium" achieved through this interaction and the "emergent smartness" of the overall environment.

3 Selected Issues and Challenges

While the area of ambient intelligence environments evidently faces a large number of issues and challenges (for a recent overview see the handbook chapter by Streitz & Privat [1]), two grand themes are highlighted here that were identified in the EU-funded project InterLink [2]: "Socially Aware Ambient Intelligence" and "Privacy, Trust, and Identity".

Social aware people are community minded and socially active in their social context. Communication between humans as part of a more comprehensive social dialog can also involve different artefacts as part of a socially aware system. Whereas embedded sensors and devices are already common in today's environments, the future challenge is the creation of intelligent or smart environments which behave in an attentive, adaptive, and active way to provide specific services in these environments. Applications and services will behave in a "socially aware" way. This means that they will provide a sense of involvement and knowledge about the social behaviour of other persons, e.g., their degree of attention, desire for customization and control, their emotional state, interests as well as their desire to engage in social interactions. Socially aware ambient environments will be composed of a collection of smart artefacts understanding social signalling and social context, resulting in the capability of improving social orientation and collective decision making.

In today's connected world, where computers mediate frequently our interaction and communication with our family, our friends and the outside world, many people suffer from the 'Big Brother' syndrome. Especially "privacy" is an elusive concept, because not everyone's sense of privacy is the same. Moreover, the expectations of privacy are unstable, because people's perception of privacy is situation-specific, or more general, context-dependent. The issues of changing views on privacy, trust and identity are mainly a result of the tricky trade-off for creating smartness. It is becoming more and more obvious that there is an interaction and balance/ trade-off between a) being able to provide intelligent support based on collecting sensor data for selecting and tailoring functionality to make the system "smart" and b) the right of people to be in control over which data are being collected, where, when, by whom, and how they are being used. This calls for the development of Privacy Enhancing Technologies (PETs) and especially for making it a standard part of system design by addressing the conflict of ubiquitous and unobtrusive data collection/provision with human control and attention in an open fashion and at an early stage of design.

4 Humane Smart Cities as an Application Umbrella

Issues and challenges identified above are still at a general level. In order to be more specific, we have to orient ourselves via demanding application scenarios. The theme of Urban Life Management in Smart Cities was selected as an umbrella scenario. Based on this, we developed the concept of "The Humane City" as our vision for the City of the Future and the future of Urban Living. In order to substantiate the relevance of this umbrella scenario, one can provide a long list of arguments. Due to space limitation, I list a selection: Already by the end of 2008, half the of the world population lived in urban areas. Thus, people speak of an Urban Age that we have entered and predict that the economic prosperity and quality of life will largely depend on the abilities of cities to reach their full potential. We can observe a development from real cities via virtual/digital cities to Hybrid Cities and then transforming them into Smart Cities. Obviously, there are many ways of addressing the challenges and issues of Hybrid and Smart Cities. One way to orient ourselves is to ask "what kind of city do we want to have? A technology-driven and technology-dominated one? Probably not! So, we developed the vision of a city where people enjoy everyday life and work, have multiple opportunities to exploit their human potential and lead a creative life. We called it "The Humane City". More details about this approach can be found in the deliverables of the InterLink project [2].

5 Proposals for a Future Research Agenda

In order to contribute to overcoming the gap between today's situation and the vision of the future as expressed in the goal "Towards the Humane City: Designing Future Urban Interaction and Communication Environments", we developed in the InterLink project 12 research lines (R1-R12) as constituents of a future research agenda. Due to space limitation, I can list here only their headings:

R1: Rationale for Humane/All-inclusive Cities (users are citizens)
R2: Tangible Interaction and Implicit vs. Explicit Interaction
R3: Hybrid Symmetric Interaction between Real and (multiple) Virtual worlds
R4: Space-Time Dispersed Interfaces
R5: Crowd and Swarm Based Interaction
R6: Spatial and Embodied Smartness (distributed cognitive systems, outside-in robot)
R7: Awareness and Feedback (sensors, physiological, environmental …)
R8: Emotion Processing (affective computing)
R9: Social Networks and Collective Intelligence
R10: Self-Organization in Socially Aware Ambient Systems
R11: Realization and User Experience of Privacy and Trust
R12: Scaling (as the major horizontal issue)

Acknowledgements

I like to thank all partners in the InterLink project and the EU for funding it, my co-chair of the InterLink Working Group "Ambient Computing and Communication Environments", Reiner Wichert, and especially the international experts participating in the four workshops that we organized and contributing to the developments of the wide range of concepts and the research agenda.

References

1. Streitz, N., Privat, G.: Ambient Intelligence. In: Stephanidis, C. (ed.) The Universal Access Handbook, pp. 60.1–60.17. CRC Press, Taylor & Francis Group (2009)
2. InterLink project, `http://interlink.ics.forth.gr`

Challenges and Limitations of Intelligent Ambient Assisted Living Environments

Reiner Wichert

Fraunhofer Ambient Assisted Living Alliance,
c/o Fraunhofer-Institut für Graphische Datenverarbeitung, Fraunhoferstr. 5,
64283 Darmstadt, Germany
Reiner.Wichert@igd.fraunhofer.de

Abstract. As a result of changing demographics, residing and being cared for in one's own familiar environment versus in an institutionalised inpatient setting is becoming the more attractive alternative for an ever increasing portion of the population. Despite its tremendous market potential, the AAL (Ambient Assisted Living) branch is still on the cusp of a mainstream breakthrough. A lack of viable business models is considered almost unanimously to be the greatest market obstacle to a broad implementation of innovative AAL systems. This paper highlights possible explanations for this deficit and shows why the AAL community has yet to arrive at joint solutions based on a unified AAL reference platform. Furthermore, this paper describes the enormous potential of AmI and AAL, as the first real opportunity for their success is provided through universAAL and AALOA.

Keywords: Ambient Assisted Living, Ambient Intelligence, market potential of AAL, business model, self-organisation, software infrastructures, AAL reference platform, standardisation, Smart Home, Fraunhofer AAL Alliance.

1 Challenges of Demographic Development

Since the early 90s, terms like "Smart Home", "Intelligent House" or "Intelligent Living" have expressed the possibility that components, devices and systems of home automation (appliances, heating and air, etc.), which normally work independently in a private household, and can communicate and cooperate when computer-supported [1]. Fundamentally, the use of resources such as energy and water is consequently minimised while the inhabitants' comfort and security is increased, as well as providing support for independent living, especially for older individuals who may be limited in their mobility, their cognitive abilities and/or motor skills. Thereby, the transition to health-related or medical applications is sometimes seamless, as various motivational or usage aspects of "intelligent" living environments often overlap. This is observable in the real life example of the automatic opening and closing of windows and blinds that – combined with the turning on and off of lighting – serve to increase the overall comfort level of inhabitants as well as increase the home's security level by reducing the chances of burglary. When further combined with proper heating and air control, home energy

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 304–309, 2010.

savings are increased, and, last but not least, the prerequisites for ill or handicapped individuals to consider or remain living in their own private surroundings are likewise being met [2].

Especially in regards to the last observation, in recent years, the exceedingly complex, highly dynamic information and communication technology domain designated “Ambient Assisted Living” (AAL) has developed into a decisive factor for the aging population across the scientific and market-oriented research landscape. One reason behind this occurrence lies in the demographic development in European societies, whereby the proportion of older individuals in the 60+ age group continues to grow. Aging is not automatically synonymous with the need for care, however, with increasing age, a greater proportion of the population relies on assistance, support and medical care. 96 % of 70+ individuals have one medical condition requiring treatment while 30% have five or more [3]. In the future, this will lead to massive costs associated with caregiving and the shrinking younger population will be even less capable of caring for the older generation: on the one hand, the financial means to cover the caregiving costs will continue to decrease and, on the other, the lack of qualified caregivers (shrinking workforce) will continue to grow. Simply put, the intergenerational pay-as-you-go pension plans in Germany and across Europe will no longer be viable in their current form. Opinion polls reveal that older individuals prefer to remain in their accustomed living environment for as long as possible, even in the eventuality of increasing reliance upon assistance and caregivers [4]. These results have recently lead to a number of research projects with prototypical singular solutions. While there is no functioning market for AAL applications at this time, marketing measures now appear to be poised on the horizon.

2 Limitations of Conventional AAL Solutions

A lack of business models as the foundation for a cooperation between ICT developers, service providers, medical device manufacturers and the housing industry is almost unanimously named as the greatest marketing hindrance to a broad implementation of innovative assisted living systems [5]. The high costs of the singular solutions, which dominate current practice, are the cause of this deficit. This very unfavourable cost-effect ratio is shown in a BMBF-sponsored study of AAL market potential and development opportunities. For cardiac insufficiency (heart failure), for example, costs of 17,300 Euros per extended year of life expectancy were determined [6]. Older individuals, however, including those in need of care, usually have not *one* medical illness or deficiency requiring compensation, but rather are confronted with *numerous, sometimes mutually exacerbating* symptoms and medical requirements [3]. Only by covering the aging person’s entire clinical picture as well as individual needs can costs really be minimised. Only then can – in the case of the person remaining longer at home - unallocated funds for extended nursing home stays (room and board, care and supervision) be included in the overall tally. Singular applications and products must become part of a total solution, albeit with considerable effort and expense. As the data exchange formats and protocols of these singular solutions are incompatible, the components of one application cannot be used by another without modification. Alterations by an expert system integrator are required to combine new and existing partial solutions, thereby making the total end solution prohibitively expensive [7]. Additionally, with singular solutions, sensors and other hardware components, as well as individual functionalities, must be installed and paid for multiple times as a matter of course, as the respective systems are only offered as complete packages. Fundamentally, current AAL solutions, such as building infrastructures outfitted with sensors and actuators, require detailed planning and installation by specialists in this field. Precisely this required level of expertise leads to the high costs currently prohibiting a successful marketing launch. The singular solutions are too expensive for private buyers and long term care insurance providers can absorb the high associated costs as little as the health insurance companies. The current economic downturn also makes mixed financing with private enterprise and state involvement more difficult. Furthermore, even if the costs were divided between the long term care insurance providers, the health insurance companies, the state, and the respective healthcare recipients themselves, the resulting high costs appear not to be sustainable by the society as a whole.

3 Platform-Based Comprehensive Concepts vs. Isolated Applications

It appears logically consistent, to forsake future monolithic approaches for complete, comprehensive solutions from one provider for an overall concept, in which various service providers and enterprises can develop individual and niche products that are easily capable of being integrated into an existing platform. Precisely this vision of AAL complete solutions was the impetus behind 13 Fraunhofer Institutes joining forces in 2007 to form the Fraunhofer Alliance Ambient Assisted Living [8]. The primary objective of the Alliance AAL was specified as the generation of a universal modular platform for AAL and Personal Health applications to be used in a variety of future projects. In order to control costs, future AAL solutions concerning the support of older persons and those requiring care must be based on a flexible platform that allows for modular expansion as to be customisable to the individual's needs, lifestyle and health progression. For interface definitions and to ensure ad-hoc interoperability at the semantic and process levels, recognised international standards or determinations, i.e., those of Continua Health Alliance or of the ongoing EU-IP project universAAL [9], whose participants include Fraunhofer Alliance AAL members, should be consulted.

The generation of an integrative platform requires a reduction in the uncontrolled growth of incompatible AAL projects, where each project develops its own new platform, similar to the IT domain, where only a relatively manageable number of predominantly standardised operating systems actually come into operation. Additional components and devices must be categorically capable of self-integrating into these infrastructures. Various EU projects that Fraunhofer Institutes of the Alliance AAL were instrumental in, have already produced validated project results with a focus on dynamic distributed infrastructures for the self-organisation of devices, sensors and services. For instance, the SOPRANO Project targets the generation of an open platform with a flexible configuration, customisable to individual needs in order to create a functioning market "ecology", in which not one lone provider develops every function and service, but rather different players work together, each bringing their own strengths to bear. In the AMIGO Project, however, a service-oriented approach and service composition techniques are utilised, such that each component implements its own integration strategies [10]. The PERSONA Project, on the other hand, is based on a bus technology in order to connect intelligent devices on an ad-hoc basis and to ensure the interoperability of devices, components and services and on the seamless integration of multimedia functions and the investigation of innovative interaction options [11].

Unfortunately, these solutions have not lead to the desired breakthrough, although each possesses substantial advantages as an AAL infrastructure. For this reason, the EU decided to sponsor only one research project within the current framework program, with the goal of developing a European platform for AAL systems. The research project selected was universAAL and it began in February 2010 with a total budget of 15 million euros and a consortium of 17 institutions from 10 countries. Within a four-year time frame, universAAL is to develop an expanded open source platform with open protocols and interfaces to ensure the interoperability of sensors, devices and services, with the groundwork for a respective standardisation being laid in stages throughout the project. The use cases and requirements have been used from the eight input projects since each project has generated suitable evaluated results. The resulting platform should serve as runtime environment for a variety of different AAL uses and applications like tracking and notification of unusual movements among elderly or to assists and motivate elderly to exercise and monitors their health. To this end, eight of the most attractive middleware platforms to emerge in the past several years, such as AMIGO, M-Power, OASIS, SOPRANO, PERSONA, and others, should be analysed and evaluated as to important criteria, such as interoperability, ease of use, dynamic expansion, integration of new components and elements, and aggregation of user interaction results and options, and, finally, with industry assistance (PHILIPS, IBM, Ericsson), be incorporated into an Ambient Assisted Living reference platform, in effect bundling the essence of these platforms.

The project strategy encompasses an array of activities for the participation of many AAL stakeholders and activists. Calls for proposals with their respective prize money should motivate scientists and enterprises worldwide to develop plug-ins to promote the platform's distribution. In various project-related workshops, experts should be consulted to quickly uncover deficits in the ongoing work in order to continually improve the platform. Via an online developer depot, developers can find or integrate specifications, development tools, sample programs, and add-ons. Additionally, the AAL community will be provided with AAL services through an online "uStore". The application options made available by universAAL will thus be enriched by numerous new contributions from the community, from which developers can also profit monetarily. For end users, the uStore becomes a kind of market hall, where they can search for suitable solutions and compare them with other users at their leisure. With the close ties of project partners DKE and Continua-Alliance to the relevant standardisation bodies, the platform's standardisation should be initiated promptly. The AAL reference platform will be available under "Open Source" licence to enable the broadest use and further development through research and industry. Fraunhofer, as the second largest project partner, acts as a driving force behind community building activities and assumes a leadership role in project-related technological activities, especially in the architectural specification of the platform and the intelligent middleware.

4 Formation of a Common AmI and AAL Community

It is of vital importance that an "ecosystem" of users and developers is established around the universAAL platform to ensure continual support and ongoing refinements. This will occur when the supporter of the universAAL platform see it as an organisational framework for their activities. Understanding that these "stalwarts" can be part of a greater AmI community, which stewards all open processes for the establishment of AmI solutions, efforts are being made through an open process involving all interested parties to found an open organisation to act as a repository for such activities. This AAL Open Association (AALOA), serving as a consortium for various projects and involving individuals as well as organisations, should form an AmI and AAL community with broad involvement from all types of stakeholders. AALOA is intended as an invitation to join in the mission of bringing together the resources, tools and people involved in AmI in a single forum, making it much easier to reach conclusions on provisions needed to design, develop, evaluate and standardise a common service platform for AmI and AAL.

A broadly distributed AAL platform unfortunately still carries with it a certain amount of risk, namely that the relatively high costs currently associated with AAL complete solutions simply cannot be shouldered by society. Therefore, business models based on a distribution of costs must be found. Not only would long term care insurance providers or health insurance companies have to participate via special AAL tariffs, but private investments would also have to take place. Unfortunately the private sector also does not wish to allocate any substantial sum of money for AAL technologies for the coming golden years. Therefore, Fraunhofer-IGD utilises the approach that this platform must be constructed such that it will not only be implemented for AAL applications, but for all AmI-based domains. The platform should consequently only mandate a few selected components and basic functionalities that are required for all AmI-related scenarios and integrate all domain-specific components as plug-ins to expand the platform in the AAL direction. By keeping a step-by-step expansion of AmI to AAL, users can than initiate their own technological investments in their younger years, for example, towards energy conservation, home automation, or leisure and games. With an open platform, the sensors, devices and services that are already installed can then be used for AAL usage without the repeat investment. This combination approach is the only way to reduce costs to a manageable level. Since AAL services and applications are only a special case of an AmI application and AmI applications will provide the templates for all other derived applications, the platform will be developed as an AmI platform which can also later be exploited by AAL. It will be easier to restrict or focus a general AmI platform for AAL than to extend a specific AAL-platform to AmI. Therefore, to ensure success, Fraunhofer Alliance AAL must work even harder to persuade the AmI community of these benefits. universAAL e.g. has allocated resources especially for community building and together with PERSONA has promised to join AALOA as one of the first projects. As a few projects and standardisation bodies have also expressed an interest in joining AALOA, it seems that we have come a significant step closer to the vision of a universal AmI platform. Parallel to working on community building, there is a clear and present need to convince the technological industry of our vision. Only when the industry is ready to develop products and services based on a common platform will AmI and AAL really get their first chance for having success.

References

1. Emiliani, P.L., Stephanidis, C.: Universal access to ambient intelligence environments: Opportunities and challenges for people with disabilities. IBM Systems Journal 44(3), 605–619 (2005)
2. Norgall, T., Wichert, R.: Ambient Assisted Living und Personal Health – Aktivitäten, Strukturen und Perspektiven der Fraunhofer-Gesellschaft; in Health Akademy, Personalisierte Medizin, Band 14 (2010)
3. Robert Koch-Institut (Hrsg.): Gesundheit im Alter. Themenheft, vol. 10. Gesundheitsberichterstattung des Bundes, Berlin (2009)
4. Wichert, R.: Configuration and Dynamic Adaptation of AAL Environments to Personal Requirements and Medical Conditions. In: HCI International 2009. Proceedings and Posters [DVD-ROM]: With 10 further Associated Conferences. LNCS, vol. 5615, pp. 267–276. Springer, Heidelberg (2009)
5. VDE Verband der Elektrotechnik Elektronik Informationstechnik e.V (Hrsg.): Positionspapier Intelligente Assistenzsysteme im Dienst für eine reife Gesellschaft., Frankfurt (2008) ISBN-978-3-8007-925512-12-4
6. Berndt, E., Wichert, R., Schulze, E., Oesterreich, D., Böhm, U., Gothe, H., Freytag, A., Daroszewska, A., Storz, P., Meyer, S., Dierks, C.: Marktpotenziale, Entwicklungschancen, Gesellschaftliche, gesundheitliche und ökonomische Effekte der zukünftigen Nutzung von Ambient Assisted Living (AAL)-Technologien. Schlussbericht des vom BMBF geförderten ITA-Projekts FKZ 16|1575 (2009), `http://publica.fraunhofer.de/`
7. Wichert, R., Norgall, T., Eichelberg, M.: AAL auf dem Weg in die Praxis – kritische Faktoren und Handlungsempfehlungen. In: Verband der Elektrotechnik Elektronik Informationstechnik: Ambient Assisted Living. Deutscher AAL-Kongress, vol. 3. VDE-Verlag, Frankfurt (2010)
8. Fraunhofer Ambient Assisted Living Alliance (2010), `http://www.aal.fraunhofer.de/`
9. universAAL: UNIVERsal open platform and reference Specification for AAL, `http://www.universaal.org/`
10. Avatangelou, E., Dommarco, R., Klein, M., Müller, S., Nielsen, C.F., Sala Soriano, P., Schmidt, A., Tazari, M., Wichert, R.: Conjoint PERSONA - SOPRANO Workshop. In: Mühlhäuser, Max (eds.) Constructing Ambient Intelligence: AmI 2007 Workshops, pp. 448–464. Springer, Heidelberg (2008)
11. Tazari, M.-R., Furfari, F., Lazaro Ramos, J.-P., Ferro, E.: The PERSONA Service Platform for AAL Spaces. In: Nakashima, H., et al. (eds.) Handbook of Ambient Intelligence and Smart Environments, pp. 1171–1199. Springer, Heidelberg (2010)

The DFKI Competence Center for Ambient Assisted Living

Jochen Frey, Christoph Stahl, Thomas Röfer,
Bernd Krieg-Brückner, and Jan Alexandersson

DFKI GmbH, Saarbrücken, Germany
{jochen.frey,christoph.stahl,thomas.roefer,bernd.krieg-brueckner}@dfki.de,
jan.alexandersson@dfki.de

Abstract. The DFKI Competence Center for Ambient Assisted Living (CCAAL) is a cross-project and cross-department virtual organization within the German Research Center for Artificial Intelligence coordinating and conducting research and development in the area of Ambient Assisted Living (AAL). Our demonstrators range from multimodal speech dialog systems to fully instrumented environments allowing the development of intelligent assistant systems, for instance an autonomous wheelchair, or the recognition and processing of everyday activities in a smart home. These innovative technologies are then tested, evaluated and demonstrated in DFKI's living labs.

Keywords: intelligent environments, ambient assisted living, living labs.

1 Vision

The DFKI Competence Center Ambient Assisted Living[1] is a cross-project and cross-department virtual organization within the German Research Center for Artificial Intelligence coordinating and conducting research and development in the area of Ambient Assisted Living. We investigate AAL from the following perspectives and underlying research areas: Cognitive Assistance (context awareness), Physical Assistance (robotics), Comfort Systems (home automation), Healthcare (telemedicine), Service Portal (multi-agent systems), and Social Connectedness (pervasive computing). Each area joins competences from one or more of the departments Intelligent User Interfaces, Safe and Secure Cognitive Systems, Agents and Simulated Reality, Institute for Information Systems, Augmented Vision, Knowledge Management and Robotics Innovation Center. Following a holistic approach to AAL, the driving force for research and development are primarily users of the technology, society and business partners. Our long-term vision is to promote an accessible intelligent environment beyond today's state-of-the-art based on standards-based open architectures and innovative solutions where everyone can continue to live a secure and autonomous life and thus play a role in society.

[1] http://ccaal.dfki.de

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 310–314, 2010.

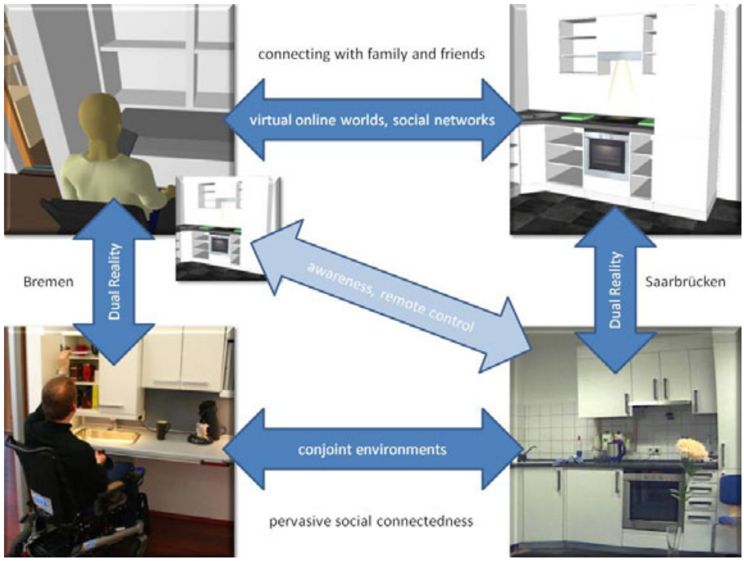


Fig. 1. The concept of Dual Reality (vertical axis) refers to the mutual influence between the real and the virtual world. Both worlds are connected to the Universal Remote Console framework thus keeping them in sync. User actions and device states in one world are reflected in the other world accordingly. The horizontal axis shows the virtual (top) and real (bottom) connections between the living labs in Bremen and Saarbruecken.

One of our current goals is to investigate social connectedness in the scope of AAL, as recent research in psychosomatic medicine indicates that social connectedness plays an important role to reduce the risk for a stroke or heart disease. Today, senior citizens are already adopting social networks such as facebook to stay in contact with their family and friends and to share fotos and news. We are interested how pervasive computing, i.e., Lifton's concept of Dual Reality [4], can further contribute to connectedness in smart homes (Figure 1). Lifton connects real and virtual environments, so that they can mutually reflect and influence each other. Similarly, Streitz and Wichert [10] put hybrid symmetric interaction on their research agenda, meaning that both worlds maintain consistency. In our living labs, we use the underlying URC technology (Section 3) to link the real environment with a virtual model thereof, so that actions performed in virtual reality (e.g., to turn on a light) have a similar effect on the physical environment. Vice versa, the virtual world reflects the real lab situation. This allows to proactively update status reports in social networks, e.g., with activity reports. Finally, we will investigate how geographically distant environments can be joined in order to create the sensation of virtual presence in conjoined environments, i.e., to let users feel like other family members were present and that they will keep an eye and know if something is wrong. As technical foundation, we extend the URC framework so that it allows us to synchronize real and

Fig. 2. Demonstrators at the BAALL: (i) remote rontrol on iPhone; (ii) instrumented flat; (iii) synchronized virtual 3-D model; (iv) autonomous wheelchair; (v) intelligent walker (photos © raumplus)

virual realities. We design and develop our AAL labs according to our Geometry-Ontology-Activity ModeL (GOAL) [9] methodology which relates activity-based requirements to a detailed 3-D model environment.

2 Ambient Assisted Living Labs

The technologies developed at DFKI are tested, evaluated and demonstrated in so called living labs. We describe two of DFKI's fully instrumented AAL environments — the Bremen Ambient Assisted Living Lab (BAALL) and the SMART KITCHEN and thereby highlight the corresponding projects and demonstrators.

2.1 Bremen Ambient Assisted Living Lab

BAALL is an automated 60 m^2 apartment suitable for the elderly and people with physical or cognitive impairments. It comprises all necessary conditions for trial living, intended for two persons, as shown in the virtual 3-D model in figure 2 (iii). BAALL aims to compensate physical impairments of the user through mobility assistance and adaptable furniture, such as a height-adaptable kitchen. Accordingly, the lab has been equipped with five automated sliding doors that open to let the wheelchair pass through when activated by a user command or proactively by the intelligent environment. Furthermore, all lights in BAALL are automated. In addition, the overall instrumentation in BAALL can be controlled remotely via mobile devices, e.g., the iPhone (Figure 2 (i)). In the EU-funded project SHARE-it, the intelligent wheelchair ROLLAND [2] and walker IWALKER [7] have been developed. They provide intelligent assistance by navigating their user within the lab by locally avoiding obstacles, controlled via the joystick or even autonomously to locations selected in a spoken natural language dialog [5].

Fig. 3. Demonstrators of the SMART KITCHEN: (i) multimodal dialogue system; (ii) activity based calendar; (iii) instrumented kitchen; (v) synchronized virtual 3-D model

2.2 Smart Kitchen

The SMART KITCHEN is a fully instrumented room that allows for realizing kitchen and living room scenarios. The SEMANTIC COOKBOOK (SC) [8] application utilizes RFID sensors to recognize products and tools on the counter and provides the user with multi-media instructions for cooking recipes. The EU-funded i2home project incorporated technical and user partners from different European countries and aimed at developing access technologies around home appliances for persons with special needs by providing the notion of pluggable user interfaces [1] and implementations thereof. The configuration of the i2home system considered here contains a wide range of appliances, e.g., televsion (Microsoft Windows Media Center), Siemens' SERVE@HOME kitchen (hood, oven, fridge, freezer, dishwasher and air condition), SweetHeart blood pressure meter, SmartLab Genie blood sugar meter, Enocean lighting, and services, e.g., Google calendar, Skype and Twitter. Figure 3(i) depicts a multimodal user interface for interaction with these appliances by speech and pointing gestures or the combination of both. A common challenge when creating intelligent user interfaces is to handle the complexity of modern appliances. The ACTIVITY BASED CALENDAR (Figure 3 (ii)) enables the users to define and schedule predefined tasks models [6]. Therefore, the users will not only be reminded of a task, the calendar automatically triggers tasks and thus assists the users by giving instructions or automating the necessary steps to perform arbitrary tasks which can also interact with the appliances connected to the system. The EU-funded VITAL project proposes a combination of advanced information and communication technologies that uses a familiar device like the television as the main vehicle for the delivery of services to elderly users in their home environments.

3 The OpenURC Alliance

One of the core concepts of the demonstrated approaches above is the usage of standards-based open architectures. We propose the Universal Remote Console (URC) framework, that focus on accessible and inclusive user interfaces by allowing any device or service to be accessed and manipulated by any controller [1]. The first project in Europe using URC technology was i2home, which had the ambitious effort to inject an ecosystem around the industrial URC standard [3] and to introduce URC technology in the field of AAL. Since i2home started, in Europe alone projects with in total 120 partners and an accumulated budget of about €80 million are currently using the URC technology, e.g., VITAL, MonAMI, Brainable, SmartSenior, SensHome etc. Therefore, we have joined an initiative to launch a global plattform for exchanging experiences, ideas and, perhaps most importantly, continue the development of the URC standard: the OpenURC Alliance. Four working groups—governance, technical, marketing and user—are shaping the organization and currently gathering partners from Europe and USA.

References

1. Frey, J., Schulz, C., Nesselrath, R., Stein, V., Alexandersson, J.: Towards Pluggable User Interfaces for People with Cognitive Disabilities. In: Proc. of the 3rd Int. Conf. on Health Informatics (HEALTHINF) (2010)
2. Krieg-Brückner, B., Röfer, T., Shi, H., Gersdorf, B.: Assistance in the Bremen Ambient Assisted Living Lab. In: Nehmer, J., Lindenberger, U., Steinhagen-Thiessen, E. (eds.) Aging and technology (special issue). GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry (to appear, 2010)
3. ISO, ISO/IEC 24752: Information Technology — User Interfaces — Universal remote console — 5 parts, International Organization for Standardization (2008)
4. Lifton, J.: Dual Reality: An Emerging Medium, Ph.D. Dissertation, Massachusetts Institute of Technology (2007)
5. Shi, H., Jian, C., Krieg-Brückner, B.: Qualitative spatial modelling of human route instructions to mobile robots. In: Proceedings of The Third International Conference on Advances in Computer-Human Interactions. IEEE Computer Society, Los Alamitos (2010)
6. Rich, C.: Building task-based user interfaces with ANSI/CEA-2018. Computer 42(8), 20–27 (2009)
7. Röfer, T., Laue, T., Gersdorf, B.: iWalker - An Intelligent Walker providing Services for the Elderly. In: Technically Assisted Rehabilitation (2009)
8. Schneider, M.: The Semantic Cookbook: Sharing Cooking Experiences in the Smart Kitchen. In: Proc. of IE 2007, Ulm. The Institution of Engineering and Technology (IET), pp. 416–423 (2009)
9. Stahl, C.: Spatial Modeling of Activity and User Assistance in Instrumented Environments, Ph.D. thesis, Saarland University (2009)
10. Streitz, N.A., Wichert, R.: InterLink Deliverable D4.2—Road-Mapping Research in Ambient Computing and Communication Environments (2009)

Intersecting the Architecture of the Internet of Things with the Future Retail Industry

Carsten Magerkurth, Stephan Haller, and Pascal Hagedorn

SAP (Switzerland) AG
Blumenbergplatz 9
CH-9000 St. Gallen
Switzerland
firstname.lastname@sap.com

Abstract. This paper discusses the approach of SAP Research in Switzerland to investigate, develop and evaluate future Internet of Things architectures and prototypes with their unique combination of three scientific pillars: SAP Research combines an environment of co-located academic education at leading universities ("Campus-Based Engineering Centers") with the concept of living laboratories in which real-world prototypes and systems are rigorously tested. SAP Research Switzerland hosts the "Future Retail Center" (FRC) in order to validate innovations in the retail industry. As an orthogonal element, we also structure our research activities in technological dimensions as opposed to the industry-specific living labs. The "Smart Items Research Program" bundles and focuses all research topics that are related to Ambient Intelligence (AmI), Internet of Things, and Pervasive Computing. With the researchers from the engineering centers, the industry focus in the living labs, and the different research projects and research programs, a holistic research perspective is created that ensures a highly effective and focused execution of research, unifying technical Internet of Things architectures with the corresponding business needs and forming a unique landscape of innovation.

1 People, Technology, Industry: The Research Landscape at SAP Switzerland

In order for research to be both innovative in the respective research discipline and to have a real-world impact or "external validity", it is crucial to regard both the core scientific questions in any given field and the demands of the related industries. The demands of the business might naturally fuel and direct the activities in a research discipline, and, likewise, technological breakthrough innovations might turn entire industries around and change the rules of the business. It is therefore necessary to acknowledge and address these interdependencies in a structural and systematic way. In order to build up effective research environments and landscapes that take both orthogonal dimensions into account, SAP Research actively tries to provide the right mix of skills and perspectives in their research personnel by following the approach of "Campus-Based Engineering Centers" (CEC) that are research labs co-located with

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 315–319, 2010.

leading universities in which PhD students participate in both the activities at the respective university and at SAP, so that they can develop both perspectives right from the start.

SAP Switzerland hosts two CECs (St. Gallen and Zürich) that are co-located with the University of St. Gallen (HSG) and the Eidgenössische Technische Hochschule Zürich (ETHZ), focusing on Business Research, Internet of Services, and the Internet of Things (which is a term we use to subsume Ambient Intelligence, Pervasvice Computing etc.). The people that conduct their research within both these universities and the SAP CEC make up the first pillar of our unique research approach. The second pillar is the technological framework that structures the research activities into several research programs across multiple CECs in close connection to the corresponding university engagement. The third pillar is a living lab (Luojus & Vilkki 2008) that brings in the industry relevant evaluation of innovations and aligns them with feedback and technologies from our partners in academia and industry.

2 The Smart Items Research Program: SAP Research's Take on Ambient Intelligence

While there are several research programs that investigate key technological areas such as User Experience or Security & Trust, the research program managed from Switzerland due to the close links to the IoT activities at ETH is called "Smart Items Research Program" and deals with anything related to the real world. Traditionally, this had mostly meant RFID and related identification technologies (Marquardt et al. 2010), but it embraces now also more AmI related topics such as integrating business logic and wireless sensor networks (Gomez et al. 2008) as well as actuation and embedded systems.

2.1 Rationale of Smart Items Research

The ongoing miniaturization of technology has resulted in a number of breakthrough technologies such as RFID, wireless sensor networks and embedded systems. While their adoption and applicability differs, they have in common that the objects in the real world augmented with such technologies now become active participants in an information technology sense. In the Smart Items research program we study the impact of these technologies on existing and newly emerging business practices by capturing and predicting the resulting automation, informational and transformational business effects. Our vision is to enable business processes to seamlessly and bi-directionally interact with real-world objects and environments on a global scale, across a variety of application domains and industries, thus realizing the Internet of Things and its benefits for business applications and services.

2.2 Challenges

Due to the successful developments over the past years, the pure technical integration of hardware technologies is playing a diminishing role; the challenges now lie in how to effectively make use of the technologies. Firstly, the use of service paradigms needs to be extended, not only for abstracting from the plethora of underlying hardware devices, but also for easily building mash-ups and integration into business applications. Services that deal directly with physical things or environments are however different compared to standard enterprise services. They are typically of lower granularity, have special constraints (e.g., locality, resources, hardware), and they might also be less reliable and not fail-safe. Secondly, more direct interaction with real-world objects and environments will enable business processes that are event-based and significantly more adaptive than today's processes. But how will these processes look like exactly in the different application areas? And even more importantly, how can they be modeled? And how to deal with the inherent unreliability of data, devices and resources of the real world?

2.3 Research Areas

Our research currently revolves around three major themes: Device integration, the Internet of Things, and novel business applications. Device integration has been in the focus over the last few years and is providing the technical foundation for the other two. The Internet of Things – as one key aspect of a Future Internet – is of significant interest internationally and seen as a game changer both in business and society, and as an enabler for new applications, processes and even business models (Haller et al. 2008). Regarding applications, the current focus is on the manufacturing, energy and retail industries, as well as the traditional areas of SCM and PLM.

3 Industry Related Evaluation in the Future Retail Center

So far, our research landscape at SAP Research Switzerland is not too different from many other labs that focus on certain research disciplines. What makes our approach unique is the consequent intersection of the scientific research programs and the validation for respective industries. At SAP Research Switzerland we have a key focus on the retail industry and try to have an impact on the respective industry solutions that SAP offers.

3.1 The Living Lab

In order to evaluate our Smart Items research, we apply our innovations to an industry relevant laboratory: The Future Retail Center (see Fig.1), one out of three Living Labs of SAP Research, is situated in Regensdorf, Switzerland and is demonstrating process innovation in various fields around retail and logistics. Historically, the Future Retail Center was founded by Migros, Siemens and SAP Business. In August 2007 it was then handed over to SAP Research Switzerland and nearly 3000 visitors came since then to have a glimpse into the future of retailing. Along the supply chain new and innovative technologies are combined on the one hand to solve old problems, and on the other hand to realize newly enabled opportunities. These demonstrators are mostly implemented in cooperation with various partners, such as

- Industry/business partners (who have the business process know-how),
- Technology partners (who have the technology know-how) and
- Academic partners (who do basic and applied research).

Fig. 1. The Future Retail Center: A Landscape for Validating Smart Items Research

The Future Retail Center offers the partners a platform for demonstrating and testing their novel technologies in a realistic and secure environment. It also allows to validate business models as well as starting discussions on standards or other open questions with experts which arise during the implementation of the demonstrators. Thus, collaborative innovation is enabled on an open platform under the management of SAP Research. Innovation Workshops are one of many possibilities to bring old and new partners together, to think about new and visionary ideas, on top of that it to prove thought leadership.

3.2 Trends and Prototypes

Current prototypes and demonstrations within the FRC relevant for AmI scenarios focus on device integration: by attempting to realize a cross-domain Internet of Things Infrastructure we connect all kinds of store related devices via a Middleware called "Real World Integration Platform" (RWIP) that integrates the point of sale (POS), weight scales, freezers etc. to the mobile phones of ordinary customers to provide real-world and real-time information from products and devices. We believe that customer-facing applications will be a game changing field especially in retail, as they will disrupt the balance of power between the various stakeholders including manufacturers and retailers. For instance, using AmI technologies to sense and assess the quality and state of goods in a store and convey this information together with personalized filtering form user profiles or other sources via a platform such as RWIP is what we see as one interesting and relevant AmI scenario for the retail domain.

4 Research Projects at the Intersection of People, Technology, and Industry

With the PhD candidates co-educated at universities and the CECs (people), the technology structure provided by the research programs (technology), and the industry focus brought in by the Future Retail Center Living Lab (industry) we address and conduct focused research projects that reflect this intersection.

One prominent example is the FP7 IP “IoT-A” (“Internet of Things-Architecture”), in which we are investigating the creation of an architectural reference model for a future Internet of Things that connects items from the real world using pervasive computing technologies. IoT-A seeks to combine the Internet of Things with the Internet of Services in order to realize business processes that involve aspects of the real world. SAP Research is driving the application of this future internet architecture to the retail domain by utilizing the FRC in order to validate an ensemble of systems and demonstrators in a real-world retail store setting.

5 Conclusions

In this paper we have outlined the three pillar approach of SAP Research involving a dual education of PhD students and researchers that consequently maps to a matrix of business industries and scientific research programs. It was discussed how the Swiss research landscape of SAP Research has a strong technical focus on Smart Items and an industry focus on retail that is reflected in the establishment of the Future Retail Center where projects such as IoT-A are conducted that accordingly bring together scientific discipline and industry validation.

References

Gomez, L., Laube, A., Sorniotti, A.: Design guidelines for integration of wireless sensor networks with enterprise systems. In: Proceedings of the 1st International Conference on Mobile Wireless Middleware, Operating Systems, and Applications, MOBILWARE, Innsbruck, Austria, February 13-15, 2008, vol. 278, pp. 1–7. ICST (Institute for Computer Sciences Social-Informatics and Telecommunications Engineering), Brussels (2007)

Haller, S., Karnouskos, S., Schroth, C.: The Internet of Things in an Enterprise Context. In: Domingue, J., Fensel, D., Traverso, P. (eds.) FIS 2008. LNCS, vol. 5468, pp. 14–28. Springer, Heidelberg (2009)

Luojus, S., Vilkki, O.: Development of user-driven research methods as the starting point for living lab activities. In: Proceedings of the 5th Nordic Conference on Human-Computer interaction: Building Bridges, Lund, Sweden, October 20-22. NordiCHI 2008, vol. 358, pp. 589–590. ACM, New York (2008)

Marquardt, N., Taylor, A.S., Villar, N., Greenberg, S.: Rethinking RFID: awareness and control for interaction with RFID systems. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems, CHI 2010, Atlanta, Georgia, USA, April 10-15, pp. 2307–2316. ACM, New York (2010)

On the Role of ExperienceLab in Professional Domain Ambient Intelligence Research

Evert van Loenen, Richard van de Sluis, Boris de Ruyter, and Emile Aarts

Philips Research Europe
HTC 34, 5656 AE Eindhoven, The Netherlands
{evert.van.loenen,richard.van.de.sluis}@philips.com,
{boris.de.ruyter,emile.aarts}@philips.com

Abstract. Concept development for professional domain AmI solutions involves different stakeholders than those for consumer products, and puts different requirements on experience test methods and facilities. Philips ExperienceLab facility for experience research is described, as well as trends and lessons learned from its use in the two domains.

Keywords: User-centered Research, Ambient Intelligence, Experience Research, Healing Environments.

1 Developing Ambient Intelligence

Ambient Intelligence (AmI) is about creating environments that are sensitive and responsive to the presence of people [1]. In 2001, the Information Society Technology Advisory Group (ISTAG) of the European Commission [2] proposed this concept to be used as the central theme for the 6th Framework Program in IST. Major technology challenges in Ambient Intelligence until today include the use of the physical world as the interface, the development of smart environments, the development of environmental programming, and eventually to master the experience itself. More recently in 2009, ISTAG issued a Report on Revising Europe's ICT Strategy [3] and developed a vision for future ICT related to sustainability, productivity, and society. For Ambient Intelligence this implies that social intelligence and design are being recognized as crucial aspects next to cognitive intelligence and computing.

Although several elements of the original ISTAG scenarios have emerged (PDA's, navigation systems [1]), in general terms the promise of Ambient Intelligence of a truly user centered technology has not yet been fulfilled over the past decade. The major reason is that most of the proposed prototypes have focused on the technology challenges, rather than identifying and solving real problems.

Finding and understanding latent needs and desires of people is a co-discovery process. For this, the so-called *Experience Research* approach was developed which identifies three levels of user-centric research: Experience@Context, Experience@Laboratory, and Experience@Field [4]. This approach is useful in separating concerns, i.e., needs from solutions and controlled from uncontrolled user studies.

Studies start *@context*, where insights are developed through contextual inquiries, observations, focus group studies and insight generation and validation methods. In

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 320–324, 2010.

subsequent ideation sessions, concepts are generated which address the needs identified. In the *@lab* phase, these concepts are developed into experience prototypes, which are presented in a realistic setting and confronted again with end-users. A range of feasibility, usability and user experience studies are then executed with participants from relevant target groups. The *@lab* studies are particularly useful for obtaining feedback from stakeholders at the earliest stage of innovation, where direct observation of the participants is desired, or where the prototype solutions are not yet compact or robust enough to use in *@field* studies.

2 AmI Development for the Consumer and Professional Domains

Like in the consumer domain, in professional domains such as retail, hospitality or healthcare, end-user experience and acceptance are crucial factors in the potential market success of a future product. However, here the end-user is *not* the person buying the product. A hotel owner for example decides on installing a particular entertainment or atmosphere solution in their hotel rooms, whereas the hotel guests are the end-users. In professional cases success or failure of a solution therefore depends not only on its attractiveness for end users, but on its attractiveness for a much larger chain of stakeholders, including owners, architects, consultants and interior designers. This has a significant impact on all phases of the concept development process, including scenario and concept demonstrator evaluations as conducted in Philips ExperienceLab.

In addition to end-user studies, methods are therefore needed to obtain input and feedback from these stakeholders in the different design phases. In the *@context* phase, stakeholder insights are built next to end-user insights. In the *@lab* phase, concept demonstrator assessments with owners, staff and other stakeholders are added.

In the next sections, the ExperienceLab is described, and its use in such multi stakeholder studies is illustrated for the case of a future shopping experience study.

3 ExperienceLab

The ExperienceLab at Philips Research Eindhoven has been designed specifically for *@lab* studies. Like the Fraunhofer inHaus [5] and Georgiatech Aware Home [6], its first instantiation was a HomeLab [7]. Over time this has grown into the ExperienceLab, which covers 3 floors and currently includes a Home, Shop, Apartment and Hospitality area. The areas themselves evolve over time, to cover new application domains. Each area houses an evolving range of concept demonstrators from a number of project teams. and is designed to enable hiding of equipment which is not part of the user experience in hollow walls and floors. Each area also has centrally controlled cameras and microphones for observation studies.

In the Home and the Apartment, a broad range of consumer product concepts have been studied, ranging from aware photo sharing to elderly support [8]. This included the Living Light concept, which resulted in the successful Ambilight television.

The Hospitality area currently consists of a hotel lobby, reception and guestroom. (See Fig. 1). Concepts which have been studied in this area range from a location aware guest inspiration system to generative art. The Shop area has been set up in the form of a complete fashion store.

Fig. 1. The lobby of the Hotel area of ExperienceLab

4 The Future Shopping Experience Case

For a study on Future Shopping Experiences, extensive @*context* studies have been undertaken, including observation studies in shopping malls and streets and participation in interview sessions with retailers in the Netherlands, London and New York. One of the observations is that shopping has become an important leisure activity: many people go shopping not because they need a specific product, but because they enjoy the activity in itself. Based on the insights, concepts aimed at helping retailers create attractive shopping experiences have been generated.

In the @*lab* phase, 8 concepts from 3 projects were prototyped and integrated in the Shop in ExperienceLab. These ranged from smart product displays which automatically adapt to the color of the accessory placed inside ("color cubes"), to an Intelligent ShopWindow, which is context aware and responds to the presence and interests of shoppers (see Fig. 2).

In addition to user studies with shoppers, stakeholder studies with retailers have been set up, to evaluate the concept demonstrators, but more importantly to use these as triggers to co-discover latent needs. For such studies a qualitative semi-structured interview format was chosen. Through close cooperation with our key account managers, representatives from major retail organizations (mainly fashion retailers) and creative specifiers (architects, retail concept designers) were invited to ExperienceLab for concept review sessions. Each concept was demonstrated in the Shop, after which open feedback was asked on its merits and drawbacks, and applicability in the stakeholders own context. Since many comments came up spontaneously during the demonstrations, the observation infrastructure of the Lab was used to record the sessions (after visitors consent) for later analysis.

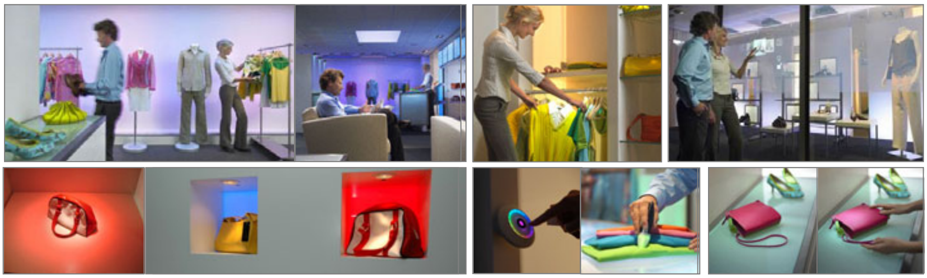


Fig. 2. Concepts integrated in the Shop. From left to right: painting with light, interactive clothing rack, intelligent shop window and color cubes, color picking and adapting spot.

This approach is markedly different from that in the end-user studies [9]. In that case an experimental design is used with participants who are non-professionals, have an adequate distribution over the target group of interest, and are properly balanced over experimental conditions for the quantitative assessments. In the retailer studies, each stakeholder already represents a specific target group, namely their customers, so understanding the reasons behind every individual comment is more important.

A number of lessons can be learned about the role of the ExperienceLab in these cases. Firstly, showing concepts as demonstrators in a realistic context was very helpful in enabling stakeholders to quickly vizualise how it would apply in their own stores, judge its potential, and identify other applications. A possible explanation may be that a demonstration takes away doubts about basic feasibility, allowing quicker focus on opportunities.

Fig. 3. Intelligent Shopwindow demonstrator. End user study condition (left) versus stakeholder study condition (right).

Secondly, the existing observation infrastructure proved quite useful for recording the stakeholder sessions. Several of the cameras had to be moved to a different position, because in the stakeholder studies the role of participant and observer is typically inverted. (See Fig. 3).

5 Outlook and Conclusions

A few trends can be observed in these AmI studies over the past years. The realization that it is crucial for the success of AmI concepts to understand and address real issues of people, has resulted in experience research methods which include @*context,* @*lab* and @*field* stages, in which increasingly not only end-users, but all stakeholders are included. This holds specifically for professional applications. Looking at future AmI opportunities, such as in healing environments in healthcare [10], this will become even more important. In those cases many stakeholders are involved (owners, specialists, nursing staff and others). In fact, the actual end-users are patients who can not be brought to a lab facility, so final evaluation of concepts has to be done in @*field* studies. However, such studies are complex and time-consuming, so should be first time right. @*lab* feasibility studies can then be executed with healthy people, before moving into the hospital.

Although the requirements for stakeholder studies are different from those for end-user studies, the same control and observation infrastructure can be used, be it with some adjustments. For both types of studies, offering realistic settings where the participants can be immersed in the envisioned experiences is crucial.

References

1. Aarts, E., Harwig, H., Schuurmans, M.: Ambient Intelligence. In: Denning, J. (ed.) The Invisible Future, pp. 235–250. McGraw Hill, New York (2001)
2. `http://cordis.europa.eu/fp7/ict/istag/home_en.html`
3. ISTAG: Revising Europe's ICT Strategy, European Commission, Luxemburg (2009), `ftp://ftp.cordis.europa.eu/pub/ist/docs/istag-revising-europes-ict-strategy-final-version_en.pdf`
4. Aarts, E.H.L., de Ruyter, B.: New research perspectives on Ambient Intelligence. Journal of Ambient Intelligence and Smart Environments 1(1), 5–14 (2009)
5. Abowd, G.: The Aware Home Research Initiative, `http://awarehome.imtc.gatech.edu/`
6. Scherer, K.: The Fraunhofer InHaus Center, `http://www.inhaus-zentrum.de`
7. de Ruyter, B., van Loenen, E., Teeven, V.: User Centered Research in ExperienceLab. In: Schiele, B., Dey, A.K., Gellersen, H., de Ruyter, B., Tscheligi, M., Wichert, R., Aarts, E., Buchmann, A. (eds.) AmI 2007. LNCS, vol. 4794, pp. 305–313. Springer, Heidelberg (2007)
8. Aarts, E., Diederiks, E.: Ambient Lifestyle, from Concept to Experience. BIS Publishers, Amsterdam (2006)
9. Kessels, A.C.J.M., Loenen, E.J., van Lashina, T.A.: Evaluating gaze and touch interaction and two feedback techniques on a large display in a shopping environment. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5726, pp. 595–607. Springer, Heidelberg (2009)
10. For example see Philips Ambient Experience, `http://www.healthcare.philips.com/in/products/ambient_experience`

The Christian Doppler Laboratory on Contextual Interfaces

Thomas Grill, Wolfgang Reitberger, Marianna Obrist,
Alexander Meschtscherjakov, and Manfred Tscheligi

HCI & Usability Unit – Christian Doppler Laboratory "Contextual Interfaces"
ICT&S Center, University of Salzburg, Sigmund Haffner G., 5020 Salzburg
{thomas.grill,wolfgang.reitberger,marianna.obrist}@sbg.ac.at,
{alexander.meschtscherjakov,manfred.tscheligi}@sbg.ac.at

Abstract. Contextual interfaces gain more and more importance within the last years. Numerous research programs that address the term context have been established, each focusing on context from their own point of view. The Christian Doppler Laboratory on Contextual Interfaces has been founded with different goals in mind. The first goal is to elaborate on the definitions and foundations of the term context to achieve a common understanding on how to use context and the terms related to context. Further we elaborate on the relation between context and user experience to tackle the complexity of the interrelations between both areas. The second goal is to apply the findings of the basic context research with two different context laboratories, each addressing a different application area. This allows to build a robust research basis that is applied to application oriented research.

Keywords: ambient interfaces, human computer interaction, context, usability.

1 Introduction

In the research area of Human-Computer Interaction (HCI) the term context gained importance within recent years. The term of context itself is primarily defined through the task context representing the parameters important to define a user interface optimized for a specific context.

In this sense it defines the possibilities to address the information necessary to be able to apply a certain functionality within our real world. Such context parameters can be named e.g. as location information and environmental parameters like noise and light conditions influencing the human during the interaction with a system.

To be able to properly address the available different aspects of context, the Christian Doppler Laboratory "Contextual Interfaces" is intensively looking into contextual ambient interaction from a qualitative viewpoint to develop a deeper understanding of optimal contextual user experiences, from a constructional viewpoint to develop improved means of interaction and interfaces for

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 325–332, 2010.

specific contexts and selected interface research areas, and from a methodological viewpoint towards detailed investigations of interaction contexts. In addition the proposed laboratory will explore the task contexts "Car" and "Factory" from these three perspectives.

2 The Research Landscape of Contextual Interfaces

The areas of research that address the topic context is widespread. When people are talking about context they address topics like the description of a design area, the environment, and the awareness about people and objects being around. Bradley et al. [2] explicit the following research areas linked to contexts.

The relation between *Computer science and Context* mainly uses three different aspects on context.

a. *Context and the application.* Moran and Dourish [4] define context as the "physical and social situation in which computational devices are embedded".
b. *Context and the user.* Dey and Abowd [3] relates context to the user's physical, social, emotional, or informational state.
c. *Context and interaction.* Schmidt et al [7] elaborated a definition of context based on a human centered interactive system where the human, the task as well as the interactive system plays an important role. Further they relate time based aspects to their context model which results in a dynamic model that evolves over time.

Within the area of ambient intelligence the term context is mainly used based on two different aspects. The first aspect of context is the *environmental context* an ambient intelligent system resides. The environment (like e.g. Smart Homes) defines the parameters and interfaces the target users interact with. It provides the playground for ambient intelligence to define, build, and evaluate intelligent systems and the user's behavior. This leads to the second aspect which is the *task context*. The task context describes the tasks an ambient intelligent system and interface shall support. It represents the basis for creating interfaces that provide optimal support for the user. In computer science frameworks addressing these contexts emerge with the intention to provide a standardized way of addressing and using contextual information within interactive systems (e.g. [1]).

This is done by a number of research institutions as well as companies. Such institutions comprise institutions like The Context-Aware Computing Groupat the Massachusetts Institute of Technology (MIT) or the Future Computing Environments (FCE) at the Georgia Institute of Technology. Companies that conduct research in the area of context are big players like Microsoft and Philips with their HomeLab where they are conducting research in ambient intelligence and on ambient user interfaces.

At our research laboratory we now regard to context as a more complex research field including perspectives on context like environmental context, social context, and the user context.

3 Contextual Interfaces at the Christian Doppler Laboratory

The Christian Doppler Laboratory (CDL) is situated at the Center for Advanced Studies and Research in Information and Communication Technologies & Society (ICT&S). The Human-Computer Interaction & Usability Unit of the ICT&S Center deals with specific questions related to design, development, and evaluation of existing and emerging interaction systems and environments. The Christian Doppler Laboratory Contextual Interfaces as part of the HCI Unit is intensively looking into contextual interaction

– from a qualitative viewpoint to develop a deeper understanding of optimal contextual user experiences,
– from a constructional viewpoint to develop improved means of interaction and interfaces for specific contexts and selected interface research areas,
– from a methodological viewpoint towards detailed investigations of interaction contexts.

In particular the proposed laboratory will explore the specific contexts "Car" and "Factory" from these three perspectives.

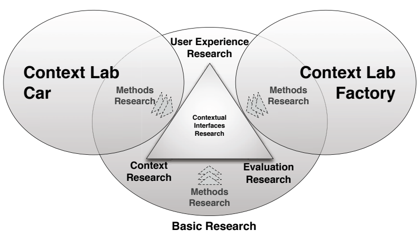


Fig. 1. Structure of the Laboratory

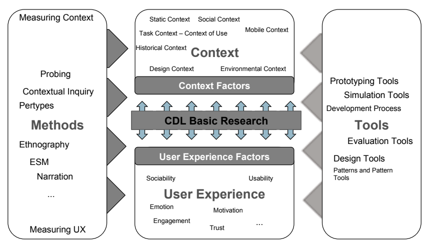


Fig. 2. Areas in the Basic Research

Figure 1 shows the basic structure of our laboratory. Our laboratory focuses on three different research areas, *Basic Research*, *Context Lab – Car*, and *Context Lab – Factory*, presented in the following subsections. The main focus of the research done within our laboratory is the area of "interface research". This means that the elaboration of user and task requirements, the design, and prototyping of highly situated interfaces and the dependency of user experience and specific context parameters reflects our main goal.

The intersections of the different areas of the laboratory show the synergies of applied research done within the *Context Lab – Car* and the *Context Lab – Factory* with the common focus on investigating in the area of contextual user interfaces through research in the different areas.

3.1 Basic Research

Basic Research at the Christian Doppler Laboratory was established with the intention to conduct basic research in the area of context and user experience with the focus on interface research.

Figure 2 shows how methods and tools applied during the research conducted in the different research areas of *Basic Research* are related. This goes inline with the two main orientations of the *Basic Research* which are defined in the vision of the Christian Doppler Laboratory. The goal to elaborate a more comprehensive User Experience and Context Model and conduct research towards methods and tools to study contextual interaction.

Context Research. Regarding context the research conducted within *Basic Research* is based on two different research topics.

a *Defining and identifying context.* The definition of context is crucial when identifying contextual parameters that shall be reflected during the requirements and design phase of contextual interfaces. A proper definition and a context model is elaborated and the validity of the context model is addressed through studies that implement contextual interfaces.

b *Designing contextual interfaces.* The second topic addressed uses the parameters identified in order to obtain prototypes of interfaces that build the basis for usability evaluations on one hand and user experience research on the other hand. The designs and the prototypes representing contextual interfaces themselves shall be developed using an applied user centered design process.

Currently a context model based on the definitions of Schmidt et al. [7], Tarasevich [8] and others is developed providing a more complete view on context from a development as well as from a user's perspective.

Evaluation Research. The goal of the evaluation research conducted in the *Basic Research* is to elaborate tools and methods to measure context parameters defined in the requirements and applied in the design of contextual interface prototypes. Measurements are addressed at different levels and fidelities of prototypes. A focus lies on the combination of different measurement and evaluation methods applied in user experience studies. Such a *multilevel measurement* approach is defined within this research area. The measurements and evaluations of the context parameters are supported through measurement and analysis tools developed.

Within this research area we establish a *Experimental Lab* that allows us to simulate context on one hand and measure contexts and context parameters on the other hand. The laboratory consists of a flexible setup of 3D projection, 3D audio, actuators and sensors, as well as a so-called "wizard of Oz" simulation working place that allows us to simulate different task as well as environmental context. The laboratory further contains observation equipment and the possibility to be linked with a usability lab residing in the same building. This allows a dynamic setup and provides us with the possibility to conduct studies and usability evaluations of prototypes within this setting. The required observation equipment is tailored to the requirements of contextual interfaces and to the observation and recording of the context parameters the evaluator is interested in.

User Experience Research. The user experience (UX) research within the *Basic Research* is multifold. On one hand a user experience model for contextual interaction is developed. This model builds the basis for the research methodology applied within the area of user experience research.

General user experience factors as well as user experience factors relevant for the specific contexts represented by the Context Labs – Factory and Car are developed. The goals behind the user experience research at the Christian Doppler Laboratory are to define UX through a clear and validated UX model. The UX Model will be operationalized into specific UX factors relevant for the two main context areas (e.g. comfort, trust, enjoyment, engagement). The UX factors are investigated through empirical studies following a grounded theory approach and further extended with literature research.

On the other hand we analyze how UX is influenced by the stakeholders of interactive processes with respect to its specified context. The UX factors identified provide important information for the interface research in the area of contextual interfaces, being the focus of our research lab.

The UX Model will be operationalized into specific UX factors relevant for the two main context areas (e.g. comfort, trust, enjoyment, engagement). The UX factors are investigated through empirical studies following a grounded theory approach and further extended with literature research.

3.2 Ambient Persuasion

The area of contextual interfaces relies especially on the use of ambient interfaces where persuasion plays an important role. We use the term *Ambient Persuasion* to refer to the use of persuasive technologies that rely on context sensing and a network infrastructure, to enable the delivery of applying context sensitive system behavior and persuasive content, personalized for the user, at the right time and at the right place.

As an example we have developed perCues, a theoretical framework for persuasive Cues in Ambient Intelligence environments. The perCues framework gives guidelines for the design of *Ambient Persuasion* applications. This goal is achieved by providing implicit and peripheral cues in the users environment to raise their awareness for certain aspects, which are relevant for their group. These cues give additional information about the users environment that would otherwise remain hidden from the user. More generally speaking, these cues act as indicators about the state of a particular group and its context [6,5]. Concrete areas of research are contextual interfaces that are designed and evaluated with the goal to avoid usage errors and thus dangerous and costly situations. In the Christian Doppler Laboratory, Ambient Persuasion is integrated as a horizontal research area, which allows to apply the theoretical foundations of this area in the two task contexts Car and Factory. Furthermore, the findings generated in the studies of Ambient Persuasion interfaces in these contexts serve as feedback to the basic research component of the lab, where they are used to further expand and refine the underlying theoretical framework. Specifically, the ongoing current research activity encompasses an operator guide for the clean room which aims to influence the operators task related decisions toward increased compliance.

3.3 Context Lab – Factory

The *Context Lab – Factory* represents one of the main research areas of the Christian Doppler Laboratory based on a cooperation with Infineon Technologies Austria AG. The main challenge for semiconductor fabrication plants therefore is the coordination of many operators working on different machines to guarantee an efficient production process. Within the *Context Lab – Factory* the context factory are thoroughly analyzed in order to optimize and introduce ambient contextual interfaces for this context. The methods developed and investigated by the *Basic Research* will be utilized and thoroughly applied throughout the studies conducted. A special focus will lie on contextual parameters identified for ambient interfaces. A selection of the main research activities for the context factory are:

- Development of a detailed context model by defining context factors influencing the user experience of operators during their working process in the factory.
- Identification of specific factory user experience factors by mapping general user experience factors on the specific factory context.
- Getting further insights on interaction approaches and interaction techniques by setting up a variety of experience studies.
- Research on setting up a *Context Lab – Factory* to execute multi-level contextual user experience investigations and measurements (e.g studying situated displays in the factory).

3.4 Context Lab – Car

Another research area of the Christian Doppler Laboratory is the *Context Lab – Car* which is based on the cooperation with AUDIOMOBIL. The automotive domain is penetrated by both well-established and upcoming technologies. Scientific effort is put into areas of the usage of advanced technology in the car from the viewpoint of users needs, preferences and limitations. Reducing the complexity of existing cars challenges the development of future cars and user features (e.g. various forms of driver assistance, night vision) by utilizing ambient intelligence and persuasive technologies. We explicitly tackle ambient persuasive systems for the context car and explore the potential of persuasion in car interfaces and develop various ambient persuasive interaction concepts through the use of ambient interfaces. The perCues framework developed in other projects provides guidelines for designing such ambient interfaces. As an example, speeding drivers can be educated towards speed limits or saving fuel by persuasive use of technology. We contribute by grounded research on advanced interaction modalities to the state of the art in the AmI field. Methods elaborated by the *Basic Research* are applied in the studies conducted in the automotive area. A selection of the main activities of the car research module are:

– Development of a thorough understanding of the context car, focusing on the usage of existing and upcoming new technologies in the car with a special focus on ambient persuasive approaches.
– Research towards a detailed identification of a context model and the definition of context factors influencing the user experience of all vehicle occupants.
– Research on car user experience (factors) based on mapping and extending a basic user experience model to the context car.
– Getting further insights on (ambient) interaction approaches and interaction techniques by setting up a variety of interface studies.
– Consolidation of collected user experience research insights in form of car user experience patterns, which do not exist for this particular context yet.

4 Conclusions

As outlined in the previous sections contextual interaction is the situated human-computer interaction and relationship dependent on a multitude of contextual factors. Our approach is to address contextual interaction from a constructional, qualitative, and methodological viewpoint in order to develop a deeper understanding of optimal contextual user experiences and its influences. Due to the focus on contextual interfaces within our research the methods and interfaces applied are strongly related to ambient intelligence and ambient interface technologies. The structure of the Christian Doppler Laboratory on "Contextual Interfaces" is set-up in a way that it allows to build a robust theoretical research basis that is applied to application oriented research.

Acknowledgements

The financial support by the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged (Christian Doppler Laboratory for "Contextual Interfaces").

References

1. Acampora, G., Loia, V., Nappi, M., Ricciardi, S.: Ambient intelligence framework for context aware adaptive applications. In: Proceedings of the Seventh International Workshop on Computer Architecture for Machine Perception, CAMP 2005, pp. 327–332. IEEE Computer Society, Washington (2005)
2. Bradley, N.A., Dunlop, M.D.: Towards a multidisciplinary user-centric design framework for context-aware applications. In: 25th -26th, pp. 25–26 (2003)
3. Dey, A.K., Abowd, G.D.: Towards a better understanding of context and context-awareness. Tech. Rep. GIT-GVU-99-22, Georgia Institute of Technology (1999)
4. Moran, T.P., Dourish, P.: Introduction to this special issue on context-aware computing. Hum.-Comput. Interact. 16(2), 87–95 (2001)

5. Obermair, C., Ploderer, B., Reitberger, W., Tscheligi, M.: Cues in the environment: a design principle for ambient intelligence. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems, CHI 2006, pp. 1157–1162. ACM, New York (2006)
6. Reitberger, W., Obermair, C., Ploderer, B., Meschtscherjakov, A., Tscheligi, M.: Enhancing the Shopping Experience with Ambient Displays: A Field Study in a Retail Store. In: Schiele, B., Dey, A.K., Gellersen, H., de Ruyter, B., Tscheligi, M., Wichert, R., Aarts, E., Buchmann, A. (eds.) AmI 2007. LNCS, vol. 4794, pp. 314–331. Springer, Heidelberg (2007)
7. Schmidt, A., Beigl, M., Gellersen, H.W.: There is more to context than location. Computers and Graphics 23(6), 893–901 (1999)
8. Tarasewich, P.: Towards a comprehensive model of context for mobile and wireless computing. In: Proc. of AMCIS 2003, pp. 114–124 (2003)

Workshop on Interaction Techniques in Real and Simulated Assistive Smart Environments

Felix Kamieth[1], Johannes Schäfer[2], Juan-Carlos Naranjo[3], Antonella Arca[4], and Jacopo Aleotti[5]

[1] Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, Germany
Felix.kamieth@igd.fraunhofer.de
[2] User Interface Design GmbH, Ludwigsburg, Germany
johannes.schaefer@uid.com
[3] ITACA - Univ. Polit. Valencia, Spain
jcnaranjo@itaca.upv.es
[4] Universidad Politécnica de Madrid, Spain
aarca@lst.tfo.upm.es
[5] Universita‘ degli Studi di Parma, Italy
aleotti@ce.unipr.it

Abstract. To further the development of AmI solutions, the simulation of solutions in virtual reality can help saving both development costs and time. The simulation of AAL environments, however, adds additional requirements to the interaction modalities used for making the simulated environment suitable for AAL target groups like elderly and handicapped people. This workshop aims at the development of a systematic approach to the creation of such simulation environments for reducing the additional gap between the virtual reality simulation of an environment and its real counterpart in terms of end-user interaction. This paper provides an analysis of the core problems of this issue an overview of current technological approaches to solve them.

Keywords: Ambient Intelligence, Virtual Reality, Simulation, Interaction.

1 Introduction

Several projects deal with realizing Ambient Intelligence scenarios using a variety of applications and devices. Within the European research community, current examples are the UNIVERSAAL project, (ICT-2009-247950), the VAALID project (IST-2007-224309), the WASP project (IST-2006-034963) and the HomeLab at the Philips Research Center in the Netherlands. Outside the European Union, examples are the Aware Home project of the Georgia Institute of Technology, the results (e.g. Intelligent Room) of the Oxygen initiative of the Massachusetts Institute of Technology and the Easy Living initiative from Microsoft Research. A challenge which is faced by all these projects is the need of connecting and integrating a multitude of different devices and services in a distributed and dynamic environment context. Furthermore, after these technical challenges have been met the challenge of interaction with these distributed and often hidden components comes into focus.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 333–336, 2010.

End users, a term that in the Ambient Intelligence field often means elderly people with no ICT background or handicapped people with special needs, need to be able to control and interact with their environment in an intuitive and user-friendly fashion. The user groups involved set high requirements in terms of usability and intuitiveness, often requiring special interaction devices, which then need to be integrated into the interaction metaphors for handling a distributed and often at least partially hidden ICT system sometimes including avatars and personal agents, immersive walls, robots and various physical devices. Important targets of research in this area are the support of different interaction styles and needs, the adaptivity of interactions to individual user preferences, end user programming of system behaviour and the important question of the balance between autonomous environment behaviour and user control.

2 Specific Requirements and Challenges of Simulated AAL Environments

With the use of simulation technology a further barrier is added between the end user and the AAL solution. Adding to the already existing technological challenge of providing senior and handicapped people with user-friendly options for interaction, the simulation environment itself requires end-user interaction. This poses a significant challenge since the already high requirements of end-users in the AAL context are increased by adding a simulation layer.

The focus of using simulation in the AAL context should be set on minimizing the additional gap introduced by the virtual reality environment. While its effects cannot be completely eliminated, increasing the immersion of the simulation and supporting appropriate interaction devices for simulation control helps the end-user in focussing on the simulated solution by minimizing the interference from the simulation itself.

3 Overview of Current Virtual Reality Solutions in the AAL Context

The use of VR applications in the AAL context is manifold. Systems are being developed to further the emotional development of autistic children [1]. Stroke rehabilitation is a further field of use – here stroke patients are kept motivated through a game-based approach [2]. Augmented reality has also been used for usability issues for elderly people, as in a windshield navigation system in [3]. The development of solutions in simulations is a further important field of research [4]. Another area of research for the use of virtual reality in service development is the improvement of access of elderly people to real environments through the use of virtual reality interfaces [5]. Virtual reality approaches are also used directly in the field of health care to offer new or improve existing services, central to which is rehabilitation work after brain damage [6]. This includes direct rehabilitation work, but also gives the opportunity for tele-rehabilitation.

4 Approaches to Increase the Immersiveness of Simulated Environments

The question of realism in virtual reality evaluations of real-world designs in the area of user interaction has been researched so far in the area of mobile devices, which are being modelled in a virtual and mixed reality setting. The aim of the most recent of these approaches [7] is to model the display and interface of mobile devices without having to build the corresponding real device and to overlay a mobile-device mock-up with the newly designed interface in a mixed-reality setting. In the area of user reactions to simulated environments and user experiences within simulations several studies have been conducted so far but the focus has been on different psychological aspects of presence within the scene like the anxiety level based on personal space invasion [8]. Regarding a sense of presence and how to enhance it, valuable insights can be gained from the current and ongoing research in the area of 3D-Games and virtual reality gaming, where a sense of presence is a requirement for a successfully developed game. Within this field of study one of the most recent papers on the subject [9] investigates the requirements for presence within Role Playing Games. Presence, as the term is used within this context, refers to “the subjective experience of being there, [10] is a psychological phenomenon that resides in the perceptions of the user”. The term presence has a wider meaning than immersion, which is considered one of the elements of presence, the others being “information intensity” and “interactivity”. Information intensity represents the detail level of information transfer from the virtual world. The simulation system sends data to the user while it receives inputs simultaneously. Interactivity means the modes of user interaction with the objects of the virtual world. The more ways there are to interact with the environment, the higher is its level of interactivity. On the subject of realism it is noted that “Scene realism is dependent on many technological factors (lighting, resolution, textures), but the consistency of the virtual environment stimuli with real-world user experiences defines the degree of mental realism, which is subjective by nature.”

5 Scientific Aim of the Workshop

The aim of this workshop is to map the needs from the AAL field in terms of interaction technology to the available technology in the VR simulation field. First, the workshop sets up an overview of current state of the art of the interaction technology available in the AAL field in a systemized fashion. The interaction technology for AAL services in the real world will then be compared with the additional specific requirements for interaction with virtual environments. Finally, an approach to systematize the simulation of AAL environments to facilitate a wide range of interaction technologies and modalities is aimed for.

This systemized approach will be the basis for future work in the area of simulating AAL environments in a way that takes the different interaction approaches already existing in the AAL field into account and thus help in the creation of simulated AAL environments from which –through end-user tests – meaningful conclusions can be drawn.

References

1. Parés, N., et al.: Promotion of creative activity in children with severe autism through visuals in an interactive multisensory environment. In: Proceedings of the 2005 Conference on Interaction Design and Children Archive, Boulder, Colorado, pp. 110–116 (2005)
2. Alankus, G., et al.: Towards Customizable Games for Stroke Rehabilitation. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems 2010, Atlanta, Georgia, USA, pp. 2113–2122 (2010)
3. Kim, S., Dey, A.K.: Simulated Augmented Reality Windshield Display as a Cognitive Mapping Aid for Elder Driver Navigation. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems 2009, Boston, MA, USA, pp. 133–142 (2009)
4. Asghar, Z.: Developing mobile services for specific user groups using Virtual environments, In: Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia 2009, article no. 17, Cambridge, United Kingdom (2009)
5. Pittarello, F., De Faveri, A.: Improving access of elderly people to real environments: a semantic based approach. In: Proceedings of the Working Conference on Advanced Visual Interfaces 2006, Venezia, Italy, pp. 364–368 (2006)
6. Flores, E.: Improving Patient Motivation in Game Development for Motor Deficit Rehabilitation. In: Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, Yokohama, Japan, pp. 381–384 (2008)
7. Steuer, J.: Defining virtual reality: Dimensions determining telepresence. Journal of Communication 42, 73–93 (1992)
8. Yang, U., Jo, D., Son, W.: Uvmode: usability verification mixed reality system for mobile devices. In: CHI 2008 Extended Abstracts on Human Factors in Computing Systems, pp. 3573–3578 (2008)
9. Nassiri, N., Powell, N., Moore, D.: Evaluation of human interactions in virtual reality environments. In: Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, New York, NY, USA, vol. 45 (2008)
10. Bostan, B.: Requirements analysis of presence: Insights from a RPG game. In: Computers in Entertainment, New York, NY, USA, vol. 7(1) (2009)

Workshop on Pervasive Computing and Cooperative Environments in a Global Context

Kirusnapillai Selvarajah and Neil Speirs

Newcastle University, UK
{K.Selvarajah,Neil.Speirs}@ncl.ac.uk

1 General Goal of the Workshop

The increasing number of devices that are invisibly embedded into our surrounding environment as well as the proliferation of wireless communication and sensing technologies are the basis for visions like ambient intelligence, ubiquitous and pervasive computing. In this context, the objective of PECES EU project is the creation of a comprehensive software layer to enable the seamless cooperation of embedded devices across various smart spaces on a global scale in a context-dependent, secure and trustworthy manner.

The purpose of the workshop is to show the problems addressed by cooperative systems and cooperative sensors in a global-scale context and show the first prototype development tools provided by the Peces project. Papers from outside the project will show complementary solutions to these problems.

The outcome of the workshop will be used to get relevant feedback on the innovation and relevance of the work performed in PECES according to the vision of external experts in Ambient Intelligence. The project will take note of the comments received and will integrate them in a public document on the assessment of the PECES first prototypes

2 Topics to Be Discussed

The key topics discussed in the workshop are the following:

- Demonstration of the solution proposed by PECES
- Middleware and frameworks for seamless cooperation through heterogeneous environments.
- Knowledge-based and context dependent systems for mobility applications

3 Preliminary Schedule

Welcome and introduction
Antonio Marqués, Project co-ordinator of PECES and Director of New technologies in ETRA

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 337–338, 2010.

PECES Middleware Challenges – On building the Bridge between Islands of Integration
Marcus Handte, Fraunhofer IAIS

PECES Application Development Tools – Simplifying the description of the context cooperating objects
Kirusnapillai Selvarajah, University of Newcastle

Selected Paper 1 (preferably) on distributed software, systems, middleware and frameworks

Selected Paper 2 (preferably) on knowledge-based systems, computational intelligence

Selected Paper 3 (preferably) on objects, devices and environments that embody ambient intelligence

Open discussion: towards the automation of the context-dependent applications development process

'Designing Ambient Interactions – Pervasive Ergonomic Interfaces for Ageing Well' (DAI'10)

Arjan Geven[1], Sebastian Prost[1], Manfred Tscheligi[1], John Soldatos[2], and Mari Feli Gonzalez[3]

[1] CURE
{geven,subasi,tscheligi}@cure.at
[2] Athens Information Technology
jsol@ait.edu.gr
[3] Fundación INGEMA
mari.gonzalez@ingema.es

Abstract. The workshop will focus on novel computer based interaction mechanisms and interfaces, which boost natural interactivity and obviate the need for conventional tedious interfaces. Such interfaces are increasingly used in ambient intelligence environments and related applications, including application boosting elderly cognitive support, cognitive rehabilitation and Ambient Assisted Living (AAL). The aim of the workshop is to provide insights on the technological underpinnings of such interfaces, along with tools and techniques for their design and evaluation.

Keywords: AAL, Elderly, User centred desing, context-awareness, pervasive computing, ergonomics.

The AmI '10 (one-day) workshop titled "*Pervasive Ergonomic Interfaces for Ageing Well*", will address the large number of emerging ergonomic and ambient interfaces that facilitate elderly inclusion and accessibility in the digital society. A large number of such interfaces and related interaction techniques (e.g., computing surfaces, voice/speech interfaces, haptics, multi-modal interfaces, context-aware interfaces, location-based interaction mechanisms) have been recently penetrating the AAL field in a number of application domain including home care, cognitive rehabilitation, social networking and interaction, experiences sharing. From a technical viewpoint the above interfaces and interaction disciplines are based on a number of emerging technologies that fall in the broader range of pervasive computing and ambient intelligence. The aim of the workshop will be to explore all aspects associated with the design, development, deployment, use and evaluation of the above-mentioned interfaces. As a result, it will cover the whole lifecycle associated with leading edge interfaces and associated interactions.

In terms of interface design the workshop will explore the effectiveness of "design-for-all" approaches, while at the same time investigating techniques for collecting user requirements, mapping them to successful designs and evaluating if they are accesible and easy to use by the elderly. From an implementation perspective, leading

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 339–341, 2010.

edge implementation technologies will be discussed, which include (but are not limited to) visual processing, acoustic processing, ergonomic devices, motivating multi-touch computing environments, emotion detection. In general, the workshop seeks proposals for novel non-obtrusive interface technologies boosting natural interactivity and facilitating task completion. Special emphasis will be paid in the way these technologies are customized to the benefit of elderly users.

As far as interface evaluation is concerned, the workshop invites ideas that pertain to all interface evaluation aspects, including impact evaluation, technical evaluation, techo-economic evaluation, deployment evaluation, evaluation with the potential final users as well as medical evaluation. Evaluation techniques will be discussed along with the experiences from evaluating real systems.

In addition to excellent research results, the workshop will attempt to present practical experiences from real systems/interfaces implementations, along with results of trials and pilot deployments. To this end, EU projects dealing with innovative interfaces and interaction techniques for ageing well and AAL applications will be invited. In general, the organizers invite researchers and practitioners working in the field of interface design, implementation and evaluation to participate in the workshop, towards a constructive sharing of ideas and experiences.

Note that the present workshop is a follow-on workshop of two other workshops held in past AmI conference, namely the AmI '09 workshop "*Designing Ambient Interactions for older users*", and the AmI'08 workshop "*Capturing AAL Needs*". Thus, the workshop takes into account the outcomes of the past workshops and is expected to significantly extend and enrich their results.

Call for Participation

Position papers are invited (but not limited to) the following topics:

- Approaches to designing elderly-friendly interfaces and related ambient interactions.
- AAL applications leveraging novel interfaces and interaction techniques / paradigms.
- Real-life interface implementations of novel ergonomic and motivating interfaces.
- Case Studies and experiences from trials or pilot operations involving interfaces for ageing well.
- Technologies underpinning the implementation novel interaction paradigm, including perceptive and multimodal interfaces.
- Innovative devices for elderly machine interactions.
- Evaluation methodologies, frameworks and experiences, including evaluations from real users.
- Pervasive context-aware systems facilitating interface implementations and/or interaction paradigms.
- Home care systems employing novel interface mechanisms for the elderly.
- Ethical and privacy questions regarding AAL systems and different interaction styles

Submissions are expected in the form of 2-4 page position papers, describing the area of research, specific work (empirical or theoretical) on the workshop topic and the innovative character of the research at hand. At least one author of accepted papers needs to register for the workshop and for one day of the conference itself. Submissions should follow the Springer LNCS guidelines.

The workshop is partially supported by the FP7 Framework Project HERMES (http://www.fp7-hermes.eu).

Important Dates:
Deadline for submissions: Friday October 1st, 2010
Feedback to authors: Friday, October 15th, 2010
Workshop DAI’10: Wednesday November 10th, 2010

3rd Workshop on Semantic Ambient Media Experience (SAME) – In Conjunction with AmI-2010

Artur Lugmayr[1], Bjoern Stockleben[2], Juha Kaario[3], Bogdan Pogorelc[4], and Thomas Risse[5]

[1] EMMi Lab., Tampere Univ. of Technology (TUT), Tampere, Finland
[2] Rundfunk Berlin Brandenburg (RBB), Berlin, Germany
[3] Varaani Oy, Tampere, Finland
[4] Jožef Stefan Institute (JSI), Ljubljana, Slovenia
[5] L3S Research Center, Hannover, Germany
lartur@acm.org, bjoern.stockleben@gmx.de, juha.kaario@varaani.fi, bogdan.pogorelc@ijs.si, risse@L3S.de

Abstract. The SAME workshop takes place for the 3rd time in 2010, and it's theme in this year was *creating the business value-creation, vision, media theories and technology for ambient media.* SAME differs from other workshops due to its interactive and creative touch and going beyond simple powerpoint presentations. Several results will be published by AMEA – the AMbient Media Association (www.ambientmediaassociation.org).

Keywords: ambient intelligence, ubiquitous computation, pervasive computation, pervasive e-commerce, media studies, ambient health.

1 Workshop Description[1]

The medium is the message! And the message was transmitted via a single distinguishable media such as television, the Web, the radio, or books. In the age of ubiquitous and pervasive computation, where the information through a distributed interlinked network of devices the question, "what is content in the age of ambient media?" becomes more and more of importance. Ambient media are embedded throughout the natural environment of the consumer – in his home, in his car, in restaurants, and on his mobile device. Predominant example services are smart wallpapers in homes, location based services, RFID based entertainment services for children, or intelligent homes. The distribution of the medium throughout the natural environment implies a paradigm change of how to think about content. Until recently, content was identified as single entities to information – a video stream, audio stream, TV broadcast. However, in the age of ambient media, the notion of content extends from the single entity thinking towards a plethora of sensor networks, smart devices,

[1] As the workshop takes place frequently, and the workshop description has been excerpted and been overtaken in parts from previous year's workshop descriptions or publications, as its content remains the same (see e.g. [2], [3] and [4]).

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 342–345, 2010.

personalized services, and media embedded in the natural environment of the user. The user actively participates and co-designs media experience with his location based input. Initiatives as the smart Web considering location based tagging for web-pages underline this development. This multidisciplinary workshop aims to answer to the challenges how to select, compose, and generate ambient content; how to present ambient content?; how to re-use ambient content and learning experiences?; what is the characteristics of ambient media, its content, and technology?; and what are ambient media in terms of story-telling and art? And finally, how do ambient media create business and value? How can ambient media be integrated into business processes and strategies? Semantics plays a crucial role in the generation of ambient media content. It can be seen as the glue between the raw data and the ambient media. Therefore we are interested to see innovative ideas how data can be (semi-automatically be interpreted and translated into media presentations. The workshop aims at a series, and at the creation of a think-tank of creative thinkers coming from technology, art, human-computer interaction, and social sciences, that are interested in glimpsing the future of semantic ambient intelligent empowered media technology.

2 Topics and Themes

The following (and related) topics are within the scope of this workshop and shall act as examples: Supply chain management with ubiquitous computation; eCommerce & ubiquitous commerce; Business processes, value-creation, and opportunities of ambient media; Understanding of the semantics of ambient content and methods for adding intelligence to daily objects; Mobile and stationary sensor data collection and interpretation algorithms and techniques; Context awareness and collection and context aware composition/selection of ambient content; Creation and maintenance of meta-information including metadata and data management; Ambient and mobile social networks, user generated content, and co-creation of content and products; Characteristics of ambient media, its content, and technological platforms; Ambient content creation techniques, asset management, and programming ambient media; Algorithms and techniques for sensor data interpretation and semantic interpretation; Applications and services, including ambient games, art and leisure content in specific contexts; Ambient interactive storytelling, narrations, and interactive advertising; Personalization, user models, multimodal interaction, smart user interfaces, and universal access; Experience design, usability, audience research, ethnography, user studies, and interface design; Business models, marketing studies, media economics, and ‘x’-commerce of semantic ambient media; Ambient interfaces (touch, gesture, haptics, biometrics); Management of information, knowledge and sapience in the context of semantic ambient media; Methods for context awareness, sensor networks, and sensor data mining; Semantic data mining and text mining for pervasive media; Semantic models, semantic interpretation for ambient media presentation; Personalization and methods for locative media.

3 Workshop Organizers

Artur Lugmayr, Tampere University of Technology (TUT) & lugYmedia Inc., FINLAND. Prof. Dr. Artur Lugmayr describes himself as a creative thinker and his scientific work is situated between art and science. Starting from 1st July 2009 he is full-professor for entertainment and media production management at the department for industrial management at the Tampere University of Technology (TUT). His vision can be expressed as to create media experiences on future emerging media technology platforms. He is the head and founder of the New AMbient MUltimedia (NAMU) research group at the Tampere University of Technology (Finland) which is part of the Finnish Academy Centre of Excellence of Signal Processing from 2006 to 2011 (http://namu.cs.tut.fi). He is holding a Dr.-Techn. degree from the Tampere University of Technology (TUT, Finland), and is currently engaged in Dr.-Arts studies at the School of Motion Pictures, TV and Production Design (UIAH, Helsinki). He chaired the ISO/IEC ad-hoc group "MPEG-21 in broadcasting"; won the NOKIA Award of 2003 with the text book "Digital interactive TV and Metadata" published by Springer-Verlag in 2004; representative of the Swan Lake Moving Image & Music Award (http://www.swan-lake-award.org/); board member of MindTrek (http://www.mindtrek.org), EU project proposal reviewer; invited key-note speaker for conferences; organizer and reviewer of several conferences; and has contributed one book chapter and written over 25 scientific publications. His passion in private life is to be a notorious digital film-maker. He is founder of the production company LugYmedia Inc. (http://www.lugy-media.tv). More about him on Google.

Björn Stockleben, RBB, GERMANY. Björn Stockleben was awarded his master's degree in Media Sciences, Media Technology and Computer Sciences from Technical University of Brunswick and Brunswick School of Arts in 2003. As a student research assistant he worked on MHP applications for the CONFLUENT and Multimedia Car Platform (MCP) project. He wrote his master thesis on ergonomic and content-specific constraints of video on mobile devices. Since April 2004 Björn Stockleben has been employed by Rundfunk Berlin-Brandenburg as project engineer for its Innovation Projects. Currently he is working on user generated content and citizen journalism for the news and youth radio departments of RBB.

Juha Kaario, Varaani Oy, FINLAND. Juha Kaario, MSc, was a Principal Member of Engineering Staff at the Nokia Research Center. He joined Nokia Mobile Phones marketing department in 1995 and moved to Nokia Research in 1997. In Nokia research Center he has worked ten years as a research manager and senior research manager for several teams including Wearable Computing (1998-2002), Personal Content (2002-2003) and Mobile Games (2003-2007). Previously he has worked in the University of Tampere (1993–94), in the Technical Research Center of Finland (1992-93) and as a co-owner in a small enterprise (1993-1996). He is one of the originators of the the Multi-User Publishing Environment (MUPE) application platform. His interest is in multi-discipline research for personal content, pervasive computing and mobile services. He is currently working with his new startup company Varaany Oy in Finland.

Bogdan Pogorelc, Ljubljana University, SLOVENIA. Bogdan Pogorelc is a Ph.D. candidate and a Research Assisstant at Department of Intelligent Systems at Jožef Stefan Institute (JSI) in Ljubljana, Slovenia. Since 2008 he has been employed at Jožef Stefan Institute and Špica International d.o.o. for which he obtained the "*Young*

Researcher" fellowship. Bogdan was visiting researcher at University "Rovira i Virgili" in Tarragona (Spain) in 2007, where he performed research on "Fuzzy Artmap neural network for assessment of metabolic syndrome". He received several awards for his research: e.g., 1st Prize at Slovenian forum of innovations for "*Intelligent security system for the surveillance of buildings"* in 2009 and awards of i) National Instruments and ii) University of Maribor for "*Mobile electrocardiograph*" in 2006. Bogdan's research interests include artificial intelligence, machine learning, (temporal-/time series-) data mining as well as applications in these areas, especially medical informatics and ambient intelligence. His main research interest is "*Behavior recognition from motion capture systems using data mining*, such as "*Automatic recognition of gait-related health problems of elderly"* or "*Automatic management of (physical) rehabilitation process for (neurologically) impaired people"*.

Thomas Risse, L3S Research Center, GERMANY. Thomas Risse works as a senior researcher at the L3S Research Center in Hannover. He received a PhD in Computer Science from the Darmstadt University of Technology, Germany in 2006. Before he joined the L3S Research Center in 2007 he lead a research group about intelligent information environments at Fraunhofer IPSI, Darmstadt. He worked in several European and industrial projects. He was the technical director of the European funded integrated project BRICKS, which aim was to build a decentralized infrastructure for distributed digital libraries. Currently he his the deputy manager of the FP7 Living Web Archive (LiWA) project. Thomas Risse's research interests are semantic evolution, data management in distributed systems, federated search, and self-organizing systems. He serves regularly as program committee member or project reviewer. He published several papers at the relevant international conferences.

References

1. Lugmayr, A.: Ambient media. Novatica 33(188), 35–39 (2007), `http://-www.ati.es/-novatica/-infonovatica_eng.html`
2. Lugmayr, A., Risse, T., Stockleben, B., Kaario, J., Laurila, K.: Special issue on semantic ambient media experiences. Multimedia Tools and Applications 44(3), 331–335 (2009), `http://-www.springerlink.com/-content/-w220782h1h508943`
3. Lugmayr, A., Stockleben, B., Risse, T., Kaario, J., Laurila, K.: ACM Multimedia 2008: 1st workshop on semantic ambient media experiences (SAME 2008) namu series Proceeding of the 16th ACM international conference on Multimedia. ACM, New York (2008)
4. Lugmayr, A., Risse, T., Stockleben, B., Laurila, K., Kaario, J.: Semantic ambient media - an introduction. Multimedia Tools and Applications (2009), `http://-www.springerlink.com/-content/-5616q82741340375`, doi:10.1007/s11042-009-0282-z

Workshop AccessibleTV
"Accessible User Interfaces for Future TV Applications"

Volker Hahn[1], Pascal Hamisu[1], Christopher Jung[1], Gregor Heinrich[2], Carlos Duarte[3], and Pat Langdon[4]

[1] Fraunhofer-IGD, Germany
[2] vsonix GmbH, Germany
[3] University of Lisbon, Portugal
[4] University of Cambridge, UK
volker.hahn@igd.fraunhofer.de, pascal.hamisu@igd.fraunhofer.de, christopher.jung@igd.fraunhofer.de, gregor.heinrich@vsonix.com, cad@di.fc.ul.pt, pml24@eng.cam.ac.uk

Abstract. Approximately half of the elderly people over 55 suffer from some type of typically mild visual, auditory, motor or cognitive impairment. For them interaction, especially with PCs and other complex devices is sometimes challenging, although accessible ICT applications could make much of a difference for their living quality. Basically they have the potential to enable or simplify participation and inclusion in their surrounding private and professional communities. However, the availability of accessible user interfaces being capable to adapt to the specific needs and requirements of users with individual impairments is very limited. Although there are a number of APIs [1, 2, 3, 4] available for various platforms that allow developers to provide accessibility features within their applications, today none of them provides features for the automatic adaptation of multimodal interfaces being capable to automatically fit the individual requirements of users with different kinds of impairments. Moreover, the provision of accessible user interfaces is still expensive and risky for application developers, as they need special experience and effort for user tests. Today many implementations simply neglect the needs of elderly people, thus locking out a large portion of their potential users. The workshop is organized as part of the dissemination activity for the European-funded project GUIDE "Gentle user interfaces for elderly people", which aims to address this situation with a comprehensive approach for the realization of multimodal user interfaces being capable to adapt to the needs of users with different kinds of mild impairments. As application platform, GUIDE will mainly target TVs and Set-Top Boxes, such as the emerging Connected-TV or WebTV platforms, as they have the potential to address the needs of the elderly users with applications such as for home automation, communication or continuing education.

Keywords: AccessibleTV, accessibility, multimodal interfaces, set-top boxes, open platforms, connected-TV, and WebTV platforms.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 346–348, 2010.

1 Objectives of the Workshop

The one-day workshop, which will be organized on November 10th in Malaga, Spain in conjunction with the AMI 2010 Conference to underline and discuss various aspects of the GUIDE project.

The European project GUIDE "Gentle user interfaces for elderly people" (www.guide-project.eu) aims to fulfil the individual accessibility needs of elderly users with mild impairments through the provision of a development toolbox for personalized, adaptive multimodal user interfaces. It targets TVs and Set-Top Boxes as application platforms for home automation, communication or continuing education addressing the needs of elderly users in the future. GUIDE will provide a set of tools that allows developers to integrate advanced accessibility features into their applications (based on [5] recommendations) with reduced development risk and costs. GUIDE will further follow an open source framework approach to ease the integration into future TV application platforms and therewith foster the dissemination of accessibility features into future applications.

The aim of the workshop is to address various requirements on software frameworks for accessible application development as the tool set envisaged in the project, as well to how TV manufacturers and TV application platform providers [6, 7] could benefit from an open framework approach as targeted in GUIDE. It should be investigated how open source can foster acceptance of such frameworks in both the device manufacturer and the developer community. The workshop will further highlight actual trends and future research challenges on adaptive multimodal interfaces.

2 Expected Contributions of the Workshop

The workshop will bring experts from industry and research specialized in accessible ICT application development, TV and set-top box manufacturers and providers of TV application platforms/middlewares [6], as well as developers of novel UI technologies.

It is mainly conceived as an invitation workshop, where experts in the field will be invited to present during panel discussions, the challenges they currently face in the integration of accessible design principles in their products, services and developments. Furthermore, participants will help in the identification of requirements on open software frameworks for accessible application development.

Panel discussions will address various aspects including but not limited to the following topics:

- State-of-the-art in inclusive design methodologies as well as multimodal interface technologies
- Requirements for open software frameworks for TV set-top boxes
- Developer requirements for a prospective framework for accessible ICT application design, given different degrees of developer experience with accessible design
- Interoperability requirements for generating standardized user models for accessible design on various platforms

- Techniques and tools that help developers of accessible interfaces and applications to validate and verify their developments in an integrated software framework. One such approach may be through the use of simulation filters and virtual users

At the end within a final session of the workshop, participants will discuss, and identify expected benefits and means of exploiting GUIDE results including aspects such as:

- The expectations and requirements for a development toolbox as the one targeted in GUIDE
- How different modalities and their combination could work towards inclusive design with the technologies available today
- The open source framework approach targeted in GUIDE for accessible interface integration
- A Handbook with lessons learnt for developers of ICT applications with accessibility requirements and other projects in the field

References

1. AT-SPI Toolkit, Assistive Technology Service Provider Interface
2. GNOME ATK, Accessibilty Toolkit
3. MSAA, Microsoft Active Accessibility API, http://msdn.microsoft.com/en-us/library/ms697707
4. JAAPI, Java Accessibility API, http://download.oracle.com/docs/cd/E17802_01/j2se/javase/technologies/accessibility/docs/jaccess-1.2/doc/core-api.html
5. CEN/CENELEC Guide 6, Guidelines for standards developers to address the needs of older persons and persons with disabilities (January 2002)
6. HbbTV, Hybrid Broadcast Broadband TV, http://www.hbbtv.org
7. OpenTV, http://www.OpenTV.com

First Workshop on Radically Innovative AAL Services

Juan-Pablo Lázaro[1] and Sergio Guillén[2]

[1] Ronda Auguste y Louis Lumiere, 23 Nave 13 46980 Paterna, Valencia
[2] ITACA – Universidad Politécnica de Valencia, Camino de Vera, SN 8G building, Access B, 3rd floor, 46022 Valencia, Spain
jplazaro@tsbtecnologias.es, sguillen@itaca.upv.es

Abstract. The following document describes the call for papers for a workshop based on identifying the key aspects that are implicit in any innovative application or service in the field of Ambient Assisted Living. Expected results will be focused in identifying the success factors that are under successful implementations and deployments of AAL Services in real conditions.

Keywords: Ambient Assisted Living, AAL Services, innovation, best practices, barriers in AAL adoption, exploitation.

1 Introduction

In the last few years, Ambient Assisted Living (AAL) paradigm has been a driver for the development of Ambient Intelligence applications specially when talking about home and mobile environments. A number of projects all over Europe have produced new applications and services for elderly people and people with special needs. This projects had the opportunity to test those AmI technologies in real trials, where innovation is faced to real life conditions and real needs. We are in the initial stages of a new potential market that still needs references and best practices to support its progressive development.

Innovation is one of the key aspects in when creating new products into a certain market. Innovation is an added value that applies not only to new ways or technologies in order to face an unsolved problem or to solve an old problem in a more effective way. Innovation has to do with creating measurable profit, as well. Customers acquire goods and services once they offer a certain value and that value has an acceptable cost by the potential customer.

Innovation is a creative task. Creativity can be promoted thanks to the application of different methodologies. In AAL, creativity will emerge from the combination of expertise about elderly needs; requirements from elderly care; and from the combination of ICT following the Ambient Intelligence paradigm.

Results of the workshop will be focused on the identification of methodologies to be innovative in the field of AAL, analysis of key factors for successful commercialization and deployment in real scenarios, identification of barriers for the spread adoption of AAL Services by the different stakeholders involved.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 349–350, 2010.

2 Topics for Discussion

Expected papers should focus on one or more of the following topics:

2.1 Innovative Ambient Assisted Living (AAL) Services

The innovation in AAL Services is understood as a new aspect or new procedure that creates an economical value to a certain amount of stakeholders involved in the development, deployment and maintenance of a certain services in the AAL domain.

Presented AAL Service(s) should demonstrate how Ambient Intelligence radically change the traditional solution and where is the benefit/profit and where the value for the different stakeholders in the AAL domain. Papers must show:

- How Ambient Intelligence is creating a value in the solution (NOTE: mandatory issue for any paper presented to this workshop).
- How stakeholders receive a clear value and/or profit from the innovative service.
- How feasible the service is: analysis of deployment and maintenance costs.
- Technical constraints.
- Boundary conditions and legal and ethical constraints for successful deployment in real life implementations.

2.2 Business Models for AAL in Europe

Particular examples for single countries or regions are acceptable as well. Present here different models and value webs that are feasible for AAL (but also for any AmI complex system).

2.3 Methodologies for Innovation

Papers addressing this topic must present different techniques and ways to find and foster innovation for AmI systems in general, but also for AAL in particular. Methodologies presented should provide real life examples that guarantee that the methodology has been proven and can be considered as a reference for improving innovation in existing or new AAL Services.

2.4 Obstacles for Development of AAL Services

Presentation of studies and analysis about identified obstacles for the wide spread of AAL solutions in real life environments and how those obstacles and barriers can be overcome.

First Workshop on Convergence and Consolidation towards Standard AAL Platform Services

Juan-Pablo Lázaro[1], Sergio Guillén[2], Babak Farshchian[3], and Marius Mikalsen[3]

[1] Ronda Auguste y Louis Lumiere, 23 Nave 13 46980 Paterna, Valencia
[2] ITACA – Universidad Politécnica de Valencia, Camino de Vera, SN 8G building, Access B, 3rd floor, 46022 Valencia, Spain
[3] SINTEF ICT. SP Andersens veg 15 B Trondheim, Norway
jplazaro@tsbtecnologias.es, sguillen@itaca.upv.es, {Babak.farshchian,Marius.mikalsen}@sintef.no

Abstract. The following document describes the call for papers for a workshop based on identifying which are the potential commonalities that are important for an AAL system, so they can be discussed and proposed for opening an standardization process. Groups of components like context-management, user interaction management or semantic description of services are frequent components and technologies that are part of an AAL system.

Keywords: Ambient Assisted Living, standardization, ontologies, platforms.

1 Introduction

In the last few years, several initiatives in the field of standardization and reference architectures for AAL system have resulted in different research initiatives that are not well aligned. They have contributed to start the research process in the field, and now it is time to consolidate those initiatives by selecting the most appropriate common aspects and by deciding which basic components can become a basis for the standardization of this types of systems. In this framework, universAAL EU funded project has emerged as the European working force trying to gather the efforts of every European partner in the field of AAL, AmI and distributed systems.

This Workshop is open to any research initiative but specially oriented to entities who have demonstrated any experience in developing and deploying AAL Services. Service providers, developers, end users' representatives... are more than welcome to submit a paper and be involved in discussion.

2 Topics for Discussion

Expected papers should focus on one or more of the following topics:

- Requirements for a platform to support AAL Services.
- Need for specialized platform components.

B. de Ruyter et al. (Eds.): AmI 2010, LNCS 6439, pp. 351–352, 2010.

- Description of layered models and services offered between layers.
- Standardization initiatives in the field of AAL. Standardization of complete systems or just parts of them.
- Missing aspects in current AAL solutions: lack of specific tools for rapid development, deployment and maintenance in order to hide the complexity of the underlying platform.

Author Index