ESPOO 2007

VTT PUBLICATIONS 629

Communications Technologies

VTT's Research Programme 2002–2006
Final Report

VTT PUBLICATIONS 629

Communications Technologies

VTT's Research Programme 2002–2006

Final Report

Markku Sipilä (ed.)

ISBN 978-951-38-6997-7 (soft back ed.)
ISSN 1235–0621 (soft back ed.)

ISBN 978-951-38-6998-4 (URL: http://www.vtt.fi/publications/index.jsp)
ISSN 1455–0849 (URL: http://www.vtt.fi/publications/index.jsp)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT, Vuorimiehentie 3, PL 2000, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 4374

VTT, Bergsmansvägen 3, PB 2000, 02044 VTT
tel. växel 020 722 111, fax 020 722 4374

VTT Technical Research Centre of Finland, Vuorimiehentie 3, P.O.Box 2000, FI-02044 VTT, Finland
phone internat. +358 20 722 111, fax + 358 20 722 4374

VTT, Tietotie 3, PL 1000, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 7013

VTT, Datavägen 3, PB 1000, 02044 VTT
tel. växel 020 722 111, fax 020 722 7013

VTT Technical Research Centre of Finland, Tietotie 3, P.O. Box 1000, FI-02044 VTT, Finland
phone internat. +358 20 722 111, fax +358 20 722 7013

Cover photograph by Tuija Soininen

Technical editing Leena Ukskoski
Text preparing Tarja Haapalainen

Edita Prima Oy, Helsinki 2007

Communications Technologies. VTT's Research Programme 2002–2006. Final Report. [Tiedonsiirtoteknologiat. VTT:n teknologiateema 2002–2006. Loppuraportti]. Ed. by Markku Sipilä. Espoo 2007. VTT Publications 629. 354 p.

Keywords telecommunication, communication technologies, wireless communication, mobile networks, autonomic networks, radio networks, cellular networks, service architecture, filter technologies, nomadic media, future homes

Abstract

The results of VTT's strategic technology theme (program) "Future Communications Technologies" are described. The theme was a research entity during the years 2002–2006 with a goal to study and develop technologies which enable users to communicate "anywhere, anytime" with the appropriate quality of service. This goal was approached on several hierarchical levels from physical electronics up to software based systems. Therefore the theme consisted of three main topical areas: networks, radio frequency technology and smart environments. At radio frequencies new kinds of microelectromechanical (MEMS) components and devices were developed, as well as selected manufacturing processes and integrated circuits. In network technology the emphasis was on the new generation mobile network with wireless multimedia applications. Basic technologies for intelligent environments were also researched and developed.

Communications Technologies. VTT's Research Programme 2002–2006. Final Report. [Tiedonsiirtoteknologiat. VTT:n teknologiateema 2002–2006. Loppuraportti]. Toim. Markku Sipilä. Espoo 2007. VTT Publications 629. 354 s.

Avainsanat Telecommunication, communication technologies, wireless communication, mobile networks, autonomic networks, radio networks, cellular networks, service architecture, filter technologies, nomadic media, future homes

Tiivistelmä

VTT:n strategisen teknologiateeman (tutkimusohjelman) "Tulevaisuuden tiedonsiirtoteknologiat" tulokset on kuvattu tässä raportissa. Teema oli vuosina 2002–2006 toteutettu tutkimuskokonaisuus. Sen päämääränä oli tutkia ja kehittää teknologioita, jotka mahdollistavat käyttäjälle riittävän laadukkaan tiedonsiirron "missä ja milloin tahansa". Päämäärää lähestyttiin useilla hierarkiatasoilla fysikaalisesta elektroniikasta ohjelmistokeskeisiin järjestelmiin. Tämän mukaisesti teema koostui kolmesta pääaihealueesta: verkot, radiotaajuusteknologia ja älykkäät ympäristöt. Radiotaajuisen teknologian alalla kehitettiin uudenlaisia mikrosähkömekaanisia (MEMS) komponentteja sekä myös valmistusprosesseja ja integroituja piirejä. Verkkoteknologiassa pääpaino oli uuden sukupolven matkaviestinverkossa ja langattomissa multimediasovelluksissa. Tutkittiin ja kehitettiin myös perusteknologioita älykkäiden ympäristöjen tarpeisiin.

Preface

The introduction of mobile communications and information networks to everyday use has changed life in a profound way during the last 20 years. The impacts can be compared to those of electricity, automobiles and the utilization of space in their own time. All of this has required an enormous research and development effort resulting in not only a large number of useful equipment but also a great body of information and new techniques in the electronics and information technology fields. VTT Technical Research Centre of Finland has participated in this work in a significant way. Because the material thus created may be of great value to science and engineering, it seemed most important to publish it as soon as possible.

A substantial part of VTT's work in the communications research field has taken place within strategic technology theme "Future Communications Technologies". This volume describes the results of that work. The task of preparing the present volume was undertaken by the staff of that theme, working in several distinct groups within VTT. I wish to express my sincere thanks to all who have contributed to this work and the theme.

Markku Sipilä

Programme Manager

Contents

1. Introduction

1.1 General

A major communications research effort at VTT Technical Research Centre of Finland took place in the strategic technology theme (programme) “Future Communications Technologies” within the timeframe 2002–2006. The results of that work are described in this volume. The present theme was one of VTT’s four strategic technology themes going on at that period of time.

1.2 Goals of the Theme

The scientific and technological goal of the theme was to study and develop technologies which enable users to communicate “anywhere and anytime” with the appropriate quality of service. This goal was approached on several hierarchical levels from physical electronics up to software based systems.

A general goal of the strategic technology themes was to facilitate a mode of operation emphasizing co-operation, networking and appropriate risk taking, so that world class results can be obtained. At the same time the applicability of the results for the industry and new business development was kept in mind.

1.3 Scientific Content Planning and Roadmapping

In the beginning, a seminar within VTT’s top information and communication technology researchers was held in February 2001, resulting in the crystallization of the following four topical areas for the theme:

- Interoperability and mobility in future networks
- Radio frequency modules
- Service architectures
- Smart human environments.

A technology roadmapping effort was then launched on each of the above areas, lasting most of 2001 and involving tens of people in various units of VTT. The goal was to clarify VTT's visions of the future, guiding the selection of research topics and helping other players to find their way in this rapidly evolving field. The result was the following publication:

M. Sipilä (ed.), "*Communications technologies: The VTT roadmaps*", VTT Tiedotteita – Research Notes 2146, Espoo, 2002.

After finishing the roadmapping effort, the actual research work took place in more than 20 projects, typically lasting from one to three years during 2002–2006. Most chapters in this Final Report describe the results of one project. Some of the projects were continuations of a previous theme project. The total amount of work in the projects of this theme was approximately 200 person-years.

The goals and research topics were under constant evaluation and evolution during the years. Consequently, at the end of the theme the above four topical areas had been transformed into the following three (Corresponding chapters in this report are given in parenthesis):

- Networks (Chapters 2–10)
- Radio frequency technology (Chapters 11–14)
- Intelligent environments (Chapters 15–22).

1.4 Management of the Theme

At the start, the guidelines and principles for all the themes were given by Professor Jarl Forstén, the Deputy Director General of VTT. A Theme Board was also formed for this theme, consisting of the following VTT persons:

- Professor Jorma Lammasniemi
- Professor Pekka Silvennoinen
- Mr Pertti Peussa.

The management of the theme was carried out by the Programme Manager, Professor Markku Sipilä, from the start of theme planning in 2000 until the end of the theme in December 2006. At the start until early 2001 he was in that position jointly with Professor Petri Mähönen.

An international Scientific Advisory Board was formed to perform the annual top expert evaluation of the theme project proposals. The members were:

- Professor José Luis Garcia-Garcia (University of Cantabria, Spain)
- Professor Iwao Sasase (Keio University, Japan)
- Professor Ramjee Prasad (Aalborg University, Denmark).

1.5 Project Selection Process

In most years an annual call for project proposals for this theme was issued within VTT. Also for ongoing multi-year projects a proposal for the next year was required. The resulting proposals were then evaluated by the Scientific Advisory Board using the following criteria:

- Scientific and technological quality
- Risk taking (encouraged)
- Are planned results world class (if achieved)?
- Business prospects or impacts.

Based on the evaluation, the Programme Manager made his recommendation on the proposals to be accepted. After comments by the Theme Board, the top management of VTT decided on the acceptance and funding of the project proposals for the next year.

1.6 Networking

A sizable amount of scientific contacts, networking and exchange of researchers was achieved within the theme and individual projects. Those fall into the following categories with respect to partners:

- Domestic (within VTT, with Finnish universities, companies etc.)
- European (strongest within EU financed projects)
- With Japan (visits to universities, research institutes and companies, joint symposia)
- With US (co-operation with e.g. Berkeley University).

1.7 Funding

The goals of the theme projects were set for obtaining results to be applied industrially a number of years in the future, therefore most of the research was done with public (Finnish government) funding. The projects were either wholly financed by VTT or jointly financed by VTT and Tekes (Finnish Funding Agency for Technology and Innovation), the European Union or companies. The funding contributions of all parties concerned are gratefully acknowledged.

2. Multimedia Over Autonomic Networks (MoNet)

Juhani Latvakoski, Eila Niemelä, Ville Ollikainen, Reijo Savola,
Chen Yan Peng, Jiehan Zhou, Teemu Väisänen, Tuomas Paaso, Juhani Laitakari,
Juha-Pekka Koivisto and Juha Kärnä

2.1 Introduction

This report summarizes the results achieved within Multimedia over Autonomic Networks [MoNet] project carried out during years 2005 and 2006 within VTT's Future Communications Technologies theme.

2.2 Goals

The context of the Multimedia over Autonomic Networks research project is visualized in Figure 1. The project aim has been to advance multimedia content management and multimedia delivery over various heterogeneous networks.

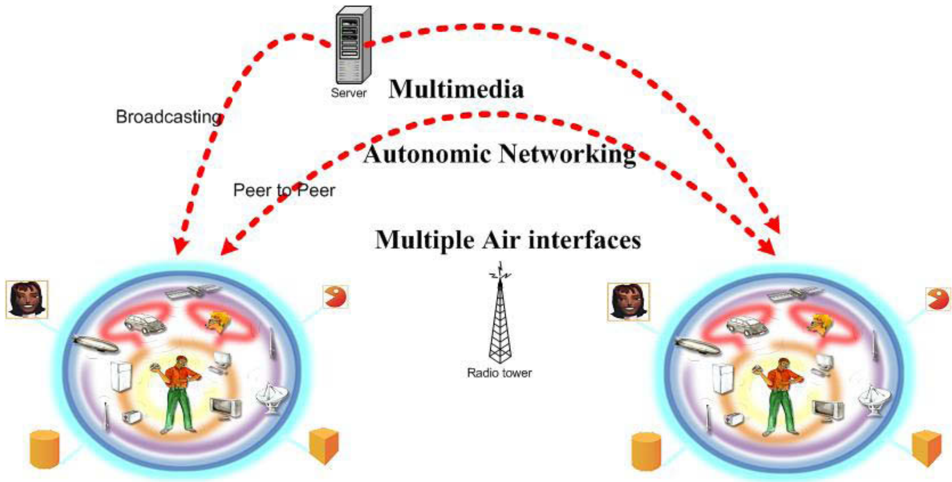


Figure 1. MoNet project.

The objectives of the project have been the following:

- To study and develop novel technical solutions for delivering multimedia over autonomic networks especially in the area of the following research items: multimedia software architectures, multimedia technologies and delivery, security and autonomic networking.
- To evaluate the technical solutions to multimedia software architectures, multimedia technologies and delivery, security and autonomic networking using separate simulations and prototypes for help.
- To integrate the selected parts into a common demonstrator to evaluate the system architecture and the provided solutions using applicable use cases with the integrated MoNet testbed.
- To disseminate the results in International scientific journals and conferences, novel patent applications, standardization contributions and doctoral theses.
- To establish novel research activities around the focus area.

2.3 Methods

In the first phase, the challenge for the project was the establishment of the common terminological and architectural view over the different kind of technologies required in enabling multimedia delivery over autonomic networks. It was a challenging task because of the heterogeneous background of the participants of the project. However, now we can say it to be very useful because different viewpoints fulfilled each others and were able to be combined to a Monet system architecture.

In the second phase, the required components for multimedia service architecture, multimedia and its delivery, autonomic networks and security was analysed, spesified, and developed. In addition, some simulations and small prototoypes were applied to validate each component separately. After this phase the results were intergrated into a common multimedia over autonomic networks demonstrator to evaluate the provided solutions using selected application use case in laboratory environment. The last phase, which is now ongoing, is dissemination of results in journals and conferences.

2.4 Results

2.4.1 Multimedia over Autonomic Networks

The results of the multimedia delivery over autonomic networks (MoNet) project is visualized in the Figure 2. The MoNet system is divided to multimedia content platform, semantic awareness, communication platform, and security. The multimedia content platform consist of content server, streaming server, and tools for content management. It incudes e.g. interfaces for adding and retrieval of content to/from the content server, means for streaming the content over DVB-H. The semantic awareness requires advanced architectures, languages, methods and tools for management of the semantic awareness information. For example, some advanced semantic awareness functions are required for finding the meaning (semantics) of the measured environmental information, which may produced by movement detector to detect movement, soapbox to measure specific issues, camera to store still images and/or video. The home server consist of application and middleware, which enables interoperation between the different technological parts of the MoNet system, and its core is communication middleware to manage the communication the other components of the system. The communication platform refers to the means for enabling end to end communication over heterogeneous networks and radio accesses. The security is seen to be the cross issue, and it is required to be embedded in all levels of the system.

The numbers in the Figure 2 refer to the enabled scenarios in the MoNet test platform executed in the laboratory environment (MoNet Lab). The (1) information content can be broadcasted via either DVB-H or 3G to a Windows terminal. The same content can be unicasted (2) via home server and the heterogeneous network to a Linux device (Nokia 770) located as a mobile node in the mobile ad hoc network. Soapbox can be used to trigger an event at home and trigger storing still images and video from home to the home server. End to end connection between home server and Linux device can be established, and the video from home can be deliverd to the Linux device over the mobile ad hoc network. The still images or video from home can be delivered as a user provided content to the content server. The user provided content can be delivered also via DVB-H broadcast channel. This kind of procedures enable e.g. real-time home surveillance system, which enables user to follow in a real-time what is happening in his/her home. This kind of a scenario is represented as a video representing the results achieved in the project.

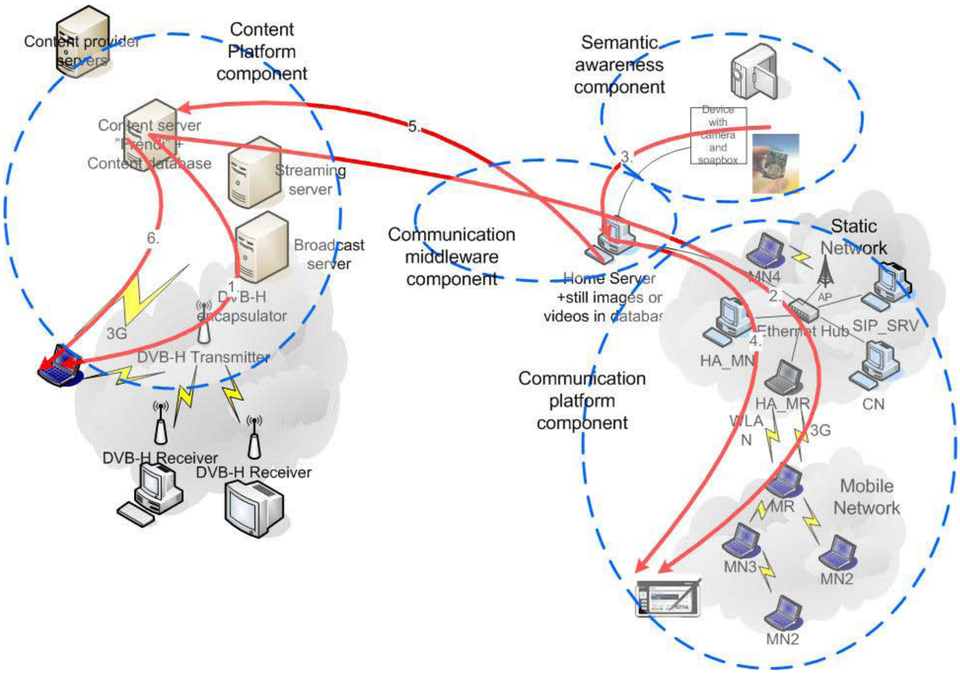


Figure 2. Multimedia delivery over Autonomic Networks.

2.4.2 Semantics Awareness

2.4.2.1 Multimedia Modelling

Semantic multimedia informatics is envisioned as an ubiquitous multimedia computational environment, in which fast-growing multimedia is easy to access and the multimedia value chain (i.e. creation, packaging, distribution and terminals) evolves from traditional vertical media (e.g. print, radio, television, storage media, telephone and Internet) dependent integration to horizontal media-independent integration [R1, R2]. In addition to the Internet, several technologies approach to semantic multimedia; 1) multimedia metamodels which enable multimedia descriptions, 2) Web multimedia languages enabling multimedia sharing over Internet, 3) Web services enabling multimedia service location, and 4) related methodologies and tools.

Enablers of Multimedia Modelling

Multimedia Service Architecture. Multimedia (i.e. audiovisual information) refers to variations of the multimedia items and contents recorded in different languages, distributed by different media, and collected by different modalities. An example multimedia service architecture (Figure 3) was defined for structuring the services required in multimedia systems [P1]. In the multimedia service architecture, the service gateway consists of middleware standards and support for representing, storing, exchanging and streaming interoperable and adaptable multimedia content over heterogeneous applications and networks. Application support services, i.e. the services used while developing multimedia applications, are classified into three categories according to their responsibilities; content management, content dissemination and content cooperation.

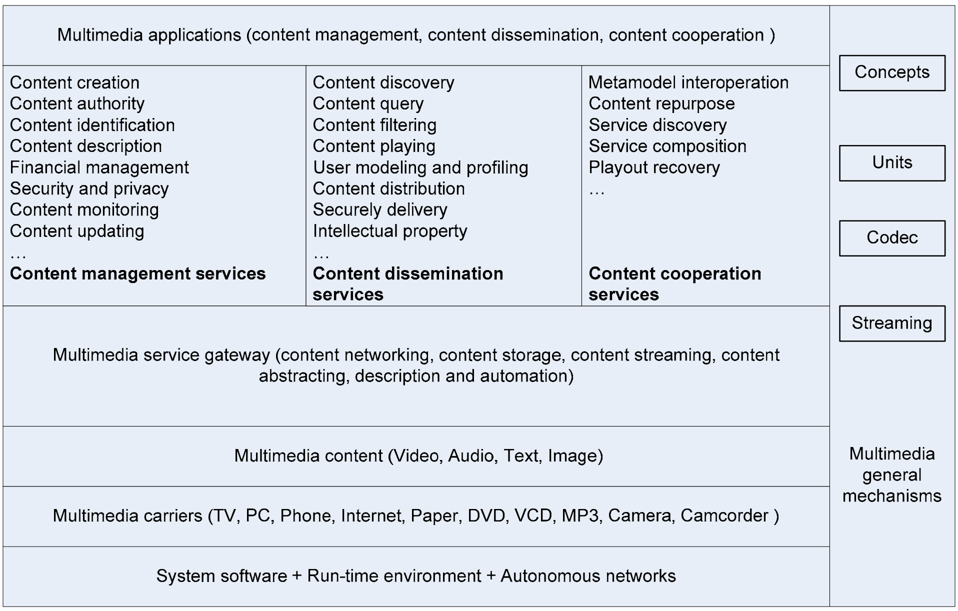


Figure 3. Conceptual multimedia service architecture.

Multimedia metamodels refer to the description of the real world with respect to three kinds of Meta information types: metadata, meta-process and meta-semantics. Metadata consists of the elements for identifying existent features of entities, e.g. CD title, CD date. Meta-process consists of the elements for specifying behavioural features of entities, e.g. workflow in engineering. Meta-semantics consists of the elements defining relational features between entities. Multimedia metamodels is structured for enhancing multimedia users to manage and develop multimedia contents across heterogeneous supporting medium (e.g. TV and networks). Similarly it consists of multimedia metadata, multimedia meta-process and multimedia meta-semantics. Although there are a number of multimedia metamodeling approaches, e.g. Digital Object Identifier (DOI) [R1], Digital Imaging, and Standard Media Exchange Framework (SMEF) [R2], a harmonized multimedia metamodels is required with support for:

- Serving different purposes (e.g. text, image, print, bibliography and metamodels exchange)
- Extensions for special interests group
- Harmonized elements with other standards and existing practices
- Extensible infrastructure for multimedia metamodels
- Support interpretation of the information's meaning from presentation to description
- Support broad range of applications.

The Web markup languages facilitate data sharing across Internet and different systems. The well-known markup languages are Hypertext Markup Language (HTML), Dublin Core (DC), extensible Markup Language (XML), Resource Description Framework (RDF) and Web Ontology Language (OWL). Especially, the OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema by providing additional vocabulary along with a formal semantics.

Multimedia Web languages combine multimedia metamodels and Web languages facilitating multimedia sharing and accessing across Internet. Generally, multimedia Web languages specify a set of multimedia metamodels elements in Web languages. There are several examples of multimedia languages; Standard Music Description Language (SMDL) in SGML [R3], P/Meta in XML [R4], Multimedia Lecture description Language MULL in XML [R5], Multimedia Retrieval Markup Language (MRML) [R6], Event Description Language (EDL) [R7], Synchronized Multimedia Integration Language (SMIL) [R8].

Methodologies for modelling semantic multimedia

Multimedia system modelling. Model-Driven Architecture (MDA) is an approach to separate the specification of system functionality from the specification of the implementation of that functionality on a specific technology platform. To this end, the MDA assists in developing a long-lived architecture that can be realized by different platform technologies, e.g. CORBA, XML, .NET [R13, R14]. In MDA, automation is much more than just automatic code generation. Automation means 1) automatic verification of software specifications to ensure that the system satisfies the requirements related to the consistence and completeness of models, 2) the crucial ability to execute abstract and incomplete models using computers, and 3) the ability to automatically generate test cases and test harnesses for automatically driving those test cases [R15, R16].

The Service-Oriented Architecture (SOA) aims at building open distributed systems that allow interaction among evolving heterogeneous devices and maximum reuse of application-independent services to increase software adaptability and efficiency. The adoption of SOA is accelerating due to the emergence of standards-based integration technologies, such as Web services. Web services based on open standards act as a facade to provide a uniform and widely accessible dynamic interface to expose business operations. XML and HTTP form the core of Web service standards. Web service technology stacks, composed of SOAP, WSDL, and OWL, are the basic building blocks that provide a solid infrastructure, enabling systems to interoperate with relative ease and reduced cost of integration. SOA allows developing software as a service delivered and consumed on demand. The benefit of this approach lies on the loose coupling of software and an increased ability to make evolving systems as, e.g., the requirements or the networked environment change.

The CommonKADS is a knowledge engineering approach to analysis, design, and management of knowledge systems. A basic characteristic of CommonKADS is the construction of model suite, where each model focuses on a limited aspect but together models provide a comprehensive view for answering three questions ‘why’, ‘what’ and ‘how’ when developing a knowledge system. In CommonKADS the organization model, the task model, the agent model, the knowledge model, the communication model and the design model are distinguished. Another major contribution of the CommonKADS is its support for structuring the knowledge model, which distinguishes three different types of knowledge. Each of them captures a related group of knowledge structures: domain knowledge, inference knowledge, and task knowledge [R9]. In summary, a multimedia informatics design methodology has to support the following:

- Taking advantage of MDA, SOA and knowledge engineering approaches
- Automation multimedia informatics system development lifecycle
- Multimedia chain – oriented metamodeling (i.e. system metamodel in the context of data, process and semantics).

Multimedia Metamodelling. Metamodelling is an activity producing metamodels. Metamodelling is closely relative to modelling, such as object-oriented modelling, which is an attempt of describing the world around us for a particular purpose [R10]. Multimedia metamodelling is viewed as a methodology and tool suites for building multimedia metamodels and multimedia base. Multimedia metamodelling is close to Object-Oriented methods and ontology engineering. There is a number of ontology development methods, e.g. Cyc method [R11], Uschold and King's method [R12], Gruninger and Fox's method [R13], KACTUS method [R14], Methontology method [R15], SENSUS-based method [R16] and On-To-Knowledge method [R17]. All of them consist of a series of interrelated activities and techniques for creating ontology. At present, there is no consensual mature and unified ontology development method [R18].

Semantic Web Services. In the example multimedia service architecture, three types of multimedia services are categorized for facilitating multimedia content management, content dissemination and content cooperation. The Semantic Web services should enable users to locate, select, employ, compose, and monitor Web-based services automatically. Web service engineering focuses on the following topics:

- Service description research addresses the mechanism and languages of documenting the exchange messages in a Web service. Several specifications are available on machine-processable service descriptions [R19].

- Automatic Web service discovery means automatic location of Web services that provide a particular service and adhere the requested constraints.
- Automatic Web service invocation involves an automatic execution of an identified Web service by a computer program or agent [R27, R29].
- Automatic Web service composition and interoperation involves an automatic selection, composition, and interoperation of Web services to perform some task, given a high-level description of an objective [R27, R29].
- In automatic Web service monitoring, individual and composite services are monitored by user requests, or the plan is changed by the actions of a software agent.
- Service annotation involves an automatic or semi-automatic annotation of service descriptions with relevant ontologies [R20].

2.4.2.2 Description of the semantic awareness demonstration

Semantic awareness component holds the context model. The sensors measure physical data from the environment which is then analyzed and used to update the contextual information in the context model. The context model is described as a semantic web, which is a graph that contains entities or classes describing the concrete elements of the environment such as room, building, computer, etc. Classes are tied to other classes with defined relations for example: “computer locates at living room” where “locates at” is the relation. Classes can also have attributes and they can extend other classes. These combined defines the graph of classes with attributes and relations between the classes.

The Semantic awareness component uses OWL Web Ontology Language to define the vocabulary which describes the relations between classes, and JENA framework to create and handle the context model. The model is created with JENA from XML file containing the ontology definitions. The advantage of using semantic description of the context information is that reasoning logic can be used to deduce new information from the model. Reasoning over the model is done with predefined logical rules which produce the deduced information from the existing model. Semantic awareness component uses a reasoner in the JENA framework for the reasoning process. Rules are predefined and they are loaded in the model on start-up.

Ontology used in semantic awareness component describes a standard home with different kind of sensor devices, Figure 4. Premises – class contains one or more Building – class instances. Building – class then contains one or more Room – classes like the normal home contains rooms. Then motion detectors can be assigned to rooms to detect if room is occupied or not. Room – class instances has an attribute to describe the current state of the room. Also Building – class has the same attribute, but its value is determined by the reasoner.

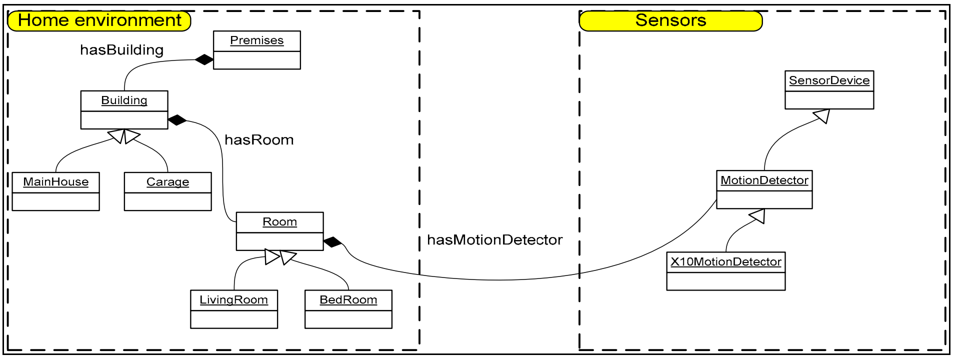


Figure 4. Description of the ontology of the home environment.

Logical rules are used in reasoning to deduce information from the contextual model. A rule consists of statements that can be chained to form the rule. Format of the statement is (subject, predicate, object). Rule has a part that is test against the context model and the triggering part which defines what is done if the first part matches the model. Format of the whole rule is [*statement*, *statement*… -> *statement*, *statement*…]. The rules used to reason if the house is occupied is listed in Figure 5. First rule determines if the house is unoccupied, second determines if one of the rooms is occupied then the house is also occupied. Third rule is to keep the model consistent by removing false statements.

```
@prefix ns: <http://anso.vtt.fi/context/ANSO-ONT.owl#>.
[
house_not_occupied:
            (?house rdf:type ns:Building) (?house ns:hasRoom ?room)
            (?room ns:hasMotionDetector ?md)
            noValue(?house ns:isOccupied)
            ->
            (?house ns:isOccupied 'false'^^xsd:boolean)
]
[
house_occupied:
            (?house rdf:type ns:Building) (?house ns:hasRoom ?room)
            (?room ns:isOccupied ?state)
            equal(?state 'true'^^xsd:boolean)
            ->
            (?house ns:isOccupied 'true'^^xsd:boolean)
]
[house_occupied_consistency:
            (?house rdf:type ns:Building) (?house ns:hasRoom ?room)
            (?room ns:isOccupied ?state)
            equal(?state 'true'^^xsd:boolean)
            (?house ns:isOccupied 'false'^^xsd:boolean)
            ->
            remove(4)
]
```

Figure 5. Reasoner rules for updating the contextual model.

To enable the context model to hold current contextual state of the house the model needs to be updated. Semantic awareness component uses X10 protocol motion detectors to detect if some room is occupied. X10 protocol motion detectors transfers the data through the electric network inside house. Special module receives the data from the electric network and transfers it to the computer where the component is running. Used X10 motion detector uses infrared to track the movement nearby and it wirelessly transfers the data to the electric network. Delay to notify a *"no movement"* state after detected motion is set to minimum which is one minute. Thus one minute period must be waited before the state changes from *"detected motion"* back to *"no detection"*.

Semantic awareness component provides an interface for applications to use the model. Model can be queried with a query language and applications can add listeners with event triggering conditions. Conditions are expressed as first order rules and when the rule matches the current information in the model notification is send to the listener. Rules used to trigger the motion detection event are shown below.

```
@prefix ns: <http://anso.vtt.fi/context/ANSO-ONT.owl#>.

Rule for motion detection event:
(ns:DigitalHome ns:isOccupied ?state) equal(?state 'true'^^xsd:boolean)

Rule for notifying when no motion is detected anymore:
(ns:DigitalHome ns:isOccupied ?state) equal(?state 'false'^^xsd:boolean)
```

Figure 6. Rules used to trigger notification events.

Component contains UPnP defined service that uses the information in the context model to notify the user on an unauthorized movement inside the house. This HouseStatus service is a standard UPnP service defined by UPnP.org. Detailed information can be found from (HouseStatus:1 Service Template v1.01". UPnP.org, URL: http://www.upnp.org/download/standardizeddcps/hvac/HouseStatus%201.0.pdf). Service defines an occupancy state and dormancy level for the home environment. Occupancy state holds the information if the house contains people and dormancy level describes if the movement is authorized or unauthorized.

HouseStatus service is build on top of the contextual model and it is using subscription – notification interface. A listener with special condition rule is registered to the component. When the rule matches the current state of the contextual model a notification is sent to the UPnP service. Service updates it state according to events received from the component.

In Semantic awareness component house's occupancy state is determined with X10 protocol motion detectors whose location is defined to the context model. With reasoning over the context model the occupancy state can be determined. If movement is detected in one of the rooms the reasoner deduces that house is occupied. On the other hand if none of the motion detectors notify movement, reasoner deduces that house is unoccupied. If no motion detectors are installed in the house occupancy state is indeterminate. House Status service's state variables are described below.

Occupancy state:

- Indeterminate – Can't be sure if the house is inhabited
- Occupied – House is inhabited
- Unoccupied – House is empty.

Dormancy level:

- Regular – House is in regular use
- Vacation – House should be empty.

Applications can query for the states via UPnP control point or register to receive notification when the state changes. Also both occupancy state and dormancy level can be set to some state by application using the UPnP service.

2.4.3 Multimedia Content Platform

The goal of this platform is to provide multimedia content to the users in the autonomous networks. The Multimedia Content Platform consists of a media content server and a video streaming server. It also provides DVB-H broadcast services via DVB-H/T network and a complementary UMTS network [P4], [P5].

2.4.3.1 Content Server

The content server is illustrated in Figure 7. The content server serves as a general content repository. The content server contains following software components:

- **Tomcat 5.0 servlet container:** Hosts web applications (i.e., Axis SOAP server and a simple client application for browsing content at the server).
- **MySQL database:** Provides persistence mechanism.
- **Scheduling application:** Triggers batch-processing tasks (i.e., delivery of video clips to the streaming server).
- **Axis SOAP Server:** Hosts Web services (is implemented as a servlet).

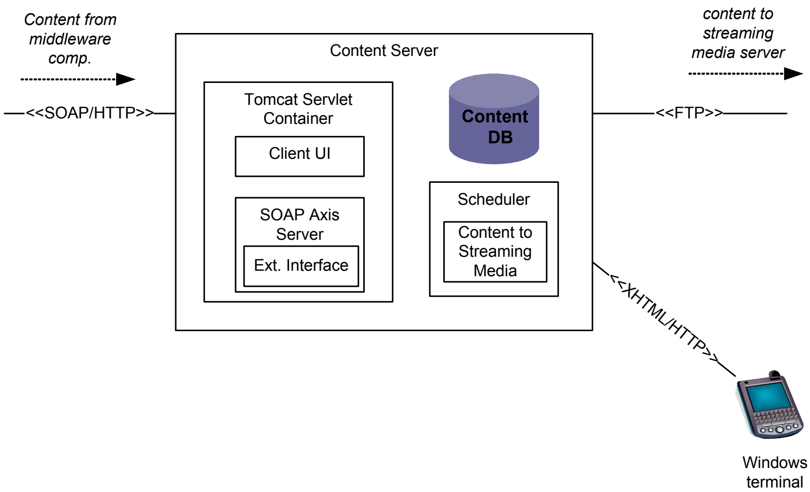


Figure 7. Content server.

The content server contains an external interface through which content can be stored to the server. The interface is implemented as a SOAP-based Web service. Soap Axis is used for building a client and server side implementation for the service. Axis provides a servlet-based implementation of the SOAP server, which will be used for hosting the Web service. Jar distributable is created for the client-side implementation of the service, so that content providers (i.e., middleware component) can access the interface with minimal effort. Figure 8 illustrates the most relevant classes for this purpose.

The video clips delivered to the server should be encoded using MPEG4, 3GPP.

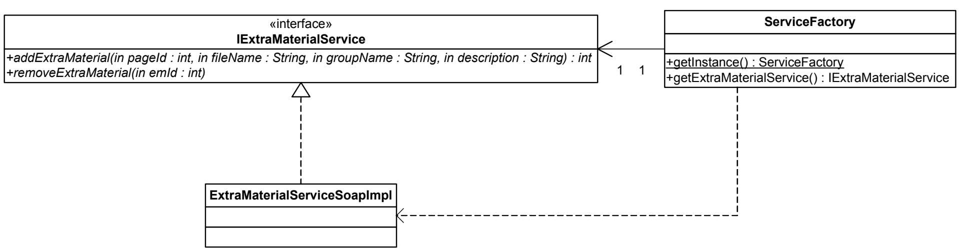


Figure 8. External Interface Class Diagram.

A simple client is implemented for browsing content at the server. The client's user interface is based on XHTML Mobile Profile 1.0. The client will be implemented as a servlet which will be deployed to the Tomcat servlet container. The initial sketch of the user interface in illustrated in Figure 9.

Figure 9. User Interface.

A batch-processing task is implemented for monitoring if there are new video clips in the server. Once arrival of a new clip is detected, it is delivered to the streaming media server via SSH. The task is triggered, e.g., once a minute.

Content server provides multimedia content such as video/audio clips (in a format 3GP, MOV, MP4, AVI, etc) and XHTML web pages. The XHTML web page files contain information, which can be browsed via both home server (in Communication middleware component) and a Linux device (in Communication platform component). The web pages are able to change on daily basis.

In a web page file, one can browse and play a video/audio clip which is embedded in the page. But the behaviour of a media player is first to download the video clip and then play it in local computer. For DVB-H broadcast services or streaming services, a video/audio clip is played in streaming way using RTP/RSTP protocols (i.e. UDP packets).

2.4.3.2 Video Streaming Server

The video streaming server within multimedia content platform component is able to stream both live video and play list which are sent from the content server to different networks. Therefore, we are able to demonstrate video/audio in two ways: download and play video clips and video streaming. In video streaming, the streaming server sends multicast video to DVB-H Encapsulator and unicast videos to both http streamer (Oulu) and DVB-H receiver (cf. Figure 10) [P4]. The challenge in this case comes from the restriction of firewalls of the networks.

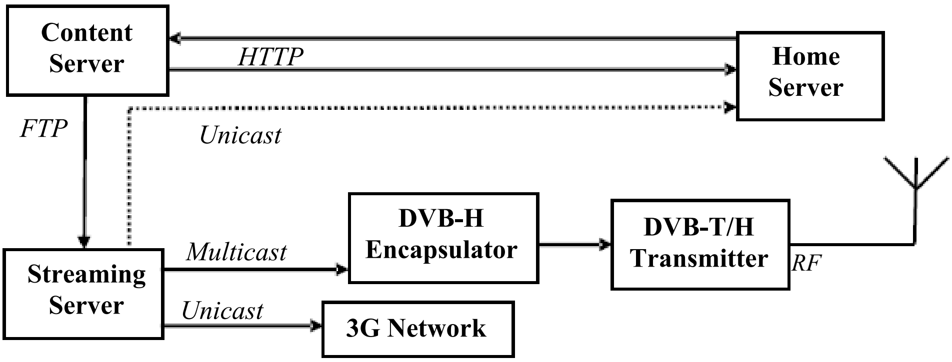


Figure 10. Content delivery diagram.

2.4.3.3 DVB-H Broadcast Multimedia Services

The nature of broadcast services is its broad coverage. Multimedia Content Platform can also deliver DVB-H broadcast multimedia services to the autonomous networks. DVB-H offers a new platform for IP-based multimedia services and contributes to universal access.

There are many challenges of providing multimedia TV experiences for DVB-H users, any where, any time. One big challenge is due to strict reception conditions and mobility. For example, when a mobile receiver enters a new cell with different transmission frequency, the receiver needs to accomplish a seamless handover process in order to continue the selected service without an interruption.

In this project, a reliable DVB-H handover approach was designed and implemented. The key method is to send one IP stream to both a DVB-H transport stream for encapsulation and UMTS network for relay without using frequency scanning (cf. Figure 12). In current implementation, the tuning failure and time-slicing synchronization can be effectively avoided.

As far as we know, the DVB-H does not have its own transmission frequency at the moment. DVB-H can be integrated into DVB-T transmission systems as DVB-H is an extension of DVB-T standard and compatible with existing DVB-T networks [P5]. We adopted the DVB-T/H network architecture of sharing multiplex with Otadigi DVB-T network. Our approach of using UMTS as alternative network saves more bandwidth for DVB-T.

2.4.3.4 Description of Content Platform Demonstration

The sample demonstration steps are listed in the following:

- Get today's multimedia content (web pages, video/audio clips) from TV or user provided content and save them in the content server.
- Deliver the content (only video files) to the video streaming server which in turn streams the video to the home server (http streamer) using unicast.
- Streaming server also sends streams to DVB-H encapsulator using multicast and to UMTS network in unicast.
- DVB-H broadcast video content on air and a DVB-H receiver can receive DVB-H service.
- When or suppose reception condition is worse, the DVB-H handover process as describes in section 4.3.2 will take place automatically and seamlessly.

2.4.4 Communication Middleware and Platform

The home server solution in the MoNet system demonstration applies the architecture of MidGate distributed service platform, see figure 11 [P6]. The architecture is divided into three layers: a communications layer, middleware layer, and service layer. In the MoNet system, the key components have been related to event notification, environment monitoring, which are described earlier together with semantic awareness, *communications middleware and autonomic network*.

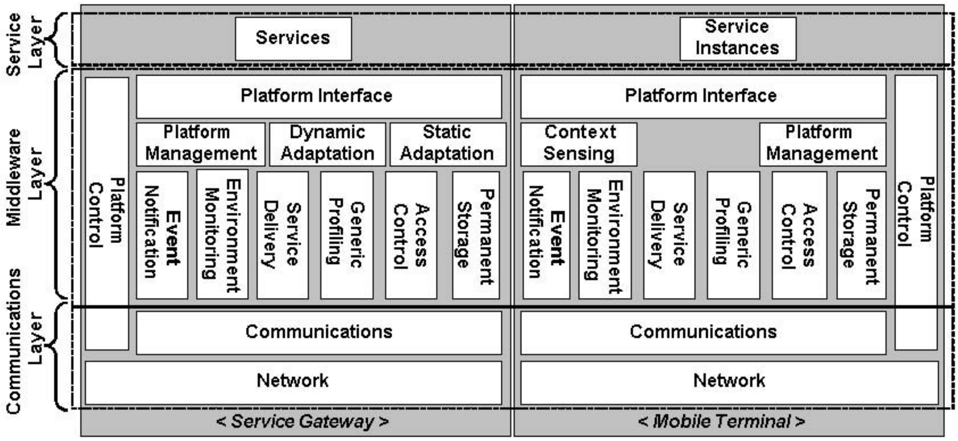


Figure 11. MidGate service platform architecture.

The MoNet autonomic network is heterogeneous mobile ad hoc network, which may or may not be connected with the static network [P7], Figure 12. A network entity can be called as mobile network, when the network changes its point of attachment to the Internet. The ad hoc network is inherently self-organized and have capability for autonomic configuration e.g. for finding the routes automatically. In the MoNet system, the key network components have been related to end to end session connectivity, network mobility, ad hoc routing and security.

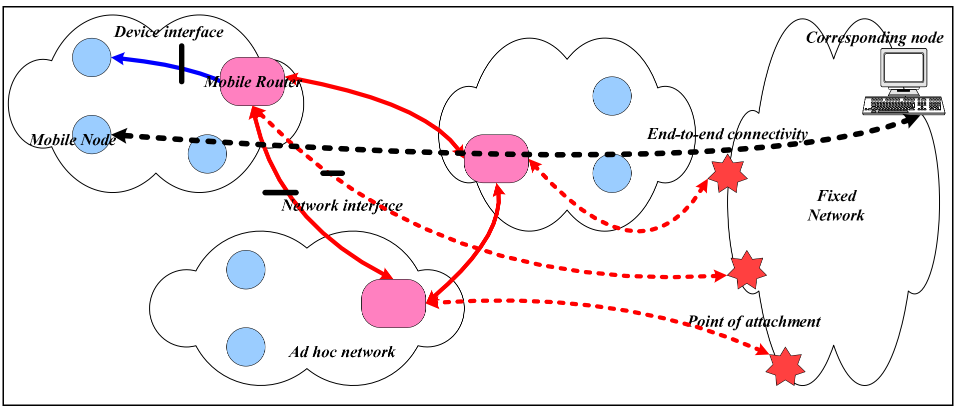


Figure 12. Heterogeneous mobile ad hoc network.

2.4.4.1 Communication Middleware Component

The Communication middleware (figure 13) consists of two separate components: a generic UPnP Control point and an application using its services. Both of the components have been implemented as OSGi bundles and are executed on the Home Server. Both of these components are implemented in Java and communication between the components will be done using Java interfaces (through OSGi).

OSGi is selected as the runtime environment as it enables to easily modify the set of available services at runtime and it also enables to easily implement new services to the platform without having to make any modifications to the middleware component itself. UPnP is selected as the implementation platform for all the components as it already provides standardized sets of services and the mechanisms for using them for most of the needed components (Camera, HouseStatus and ControlPoint).

The UPnP Application using the UPnP Devices through the ControlPoint is implemented as an OSGi bundle and it acts as “the brains” of the middleware component. It is responsible for monitoring the HouseStatus, users’ presence-state, using the camera, notifying the user of alarms in the house, inserting data into the Content Platform database and getting data from the same database to be presented to users using the Tablet Device.

Three types of UPnPDevices are needed for the demonstration: MWDevice (HouseStatus service), DigitalSecurityCamera and the TabletDevice. The HouseStatus service and the CameraDevice with all its services are standardized by the UPnP community (upnp.org) and will also be running on the same OSGi platform with the generic ControlPoint and the UPnPApplication. The TableDevice is used for monitoring the user’s online status, notifying the user of changes in the HouseStatus and providing the user the means for viewing video and pictures from the security camera.

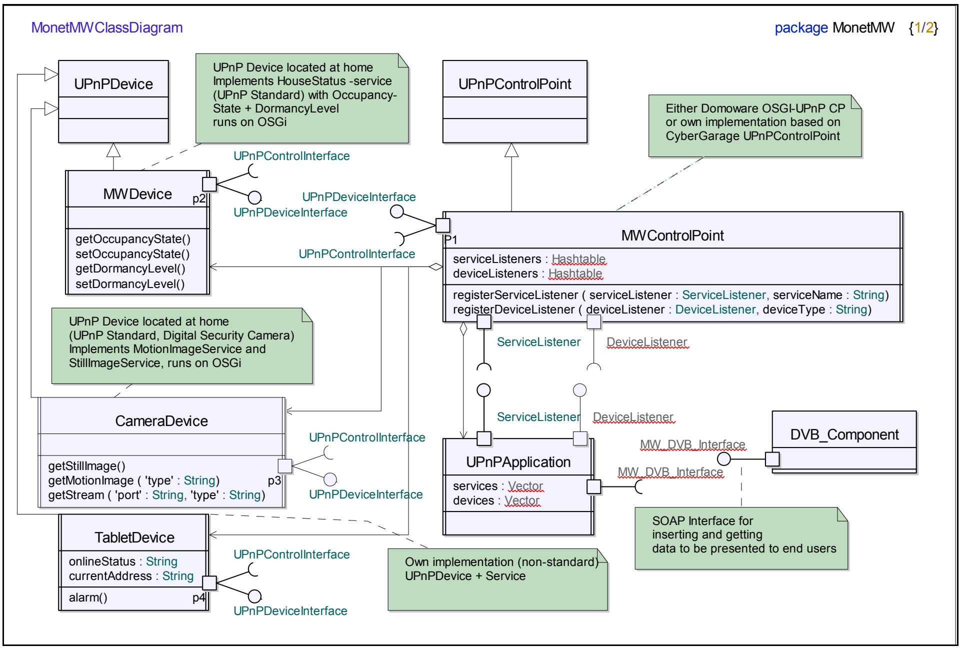


Figure 13. Communication middleware class diagram.

The home server includes UPnPControlPoint running on OSGi and MonetApp that uses the OSGI UPnP ServiceDiscovery component through OSGi. Camera and ContextMonitoringComponent also run on OSGi but they are UPnPServices, and so communication is done through UPnP. All services, Camera, ContextMonitoringComponent and Tablet, are wrapped as Monet Middleware services and can be used by MonetAPP without knowledge of UPnP. Applications can register to the OSGI Control point as listeners to services, devices and Events with certain conditions to receive notifications of these and to use them.

2.4.4.2 Description of Autonomic Network Demonstration

The applied solutions in the MoNet demonstrator have been based on session initiation protocol (SIP), IPv6, mobile IPv6, network mobility (NEMO) extensions for IPv6 and Simplified Ad-hoc On-demand Distance Vector (sAODV). In the solution, the mobile ad hoc network is connected to the static IPv6 Internet by the Mobile Router (MR). The MR executes NEMO Basic Support Protocol, which enables transparent mobility for the ad hoc network. In it, all traffic between the mobile network and the Internet goes through bi-directional tunnel, which is created between MR and Home Agent (HA). Mobile Routers operate by changing their mobility bindings according to the location of the Mobile Router. When the HA receives a packet destined to the node in the mobile network, it is encapsulated and transferred through the tunnel to the MR, which decapsulates the packet and delivers it to the final destination. NEMO protocol is transparent to the mobile nodes in the ad hoc network, and therefore the session continuity is provided for the nodes behind it in the ad hoc network.

When the mobile router moves away from the reach of WLAN, traffic is routed through 3G connection. This so called vertical NEMO handover is triggered by the router advertisements that are sent by HA_MR, via different network interfaces; 'Home Interface' and 'Visitor Interface'. When MR notices the router advertisement sent by 'Visitor Interface', it starts the address autoconfiguration process on the same interface that it got the router advertisement, and sends a BU message to HA_MR, thus establishing the bidirectional IPv6 tunnel between them. The 3G-WLAN handover happens automatically transparently to the end point of the communication, Nokia 770, which is in the role of a mobile node. This means that the end to end session continues and user see only a shorter or longer break in the video when the handover occurs, and the video continues after the break.

2.4.5 Security

In the future telecommunication systems it is very important to offer enough security for applications depending on their criticality and the context. The correct level of security functionality in the systems is typically a trade-off between security, usability and cost effectiveness. The required security level, or performance, of a system depends on particular security requirements of the system, its use cases, the context and the estimation of the impact of threats. Different applications results typically in different security solutions. In order to be able to quantify the security performance, it is important to develop methods and techniques to gather evidence of the information security level, both from technical systems (security evaluation) and from life-cycle management and information security management practices of the vendors (auditing).

2.4.5.1 Gathering Evidence of Information Security Level – Holistic Perspective by Evaluation and Auditing

Information security evaluation of software-intensive systems typically relies heavily on the experience of the security professionals. Obviously, automated approaches are needed in this field. Unfortunately, there is no practical approach to carrying out security evaluation in a systematic way. We introduce a general-level holistic framework for security evaluation and auditing based on security behaviour modelling and security evidence collection, and discusses its applicability to the design of security evaluation experimentation set-ups in real-world systems.

Products and services, and the technical infrastructures that enable them are showing a strong trend towards convergence and networking. At the same time, industrial companies and other organizations are creating very complex value nets to design and manufacture products and to maintain them. They must co-operate with their partners, subcontractors and clients, and obey regulations laid down by different governmental bodies. Figure 14 depicts this complex division of control perimeters. These trends, together with pressure from information security and privacy legislation, are increasing the need for adequately tested and managed information security solutions in software intensive systems and networks. The lack of appropriate information security solutions might have serious consequences for business and the stakeholders. Furthermore external parties demand management actions to be taken for securing information security, for example Sarbanes Oxley Act a.k.a SOX and Basel II are examples of such demands from regulators ([R22], [R23], [R27], [R29], see also [R24], [R30], [R25]).

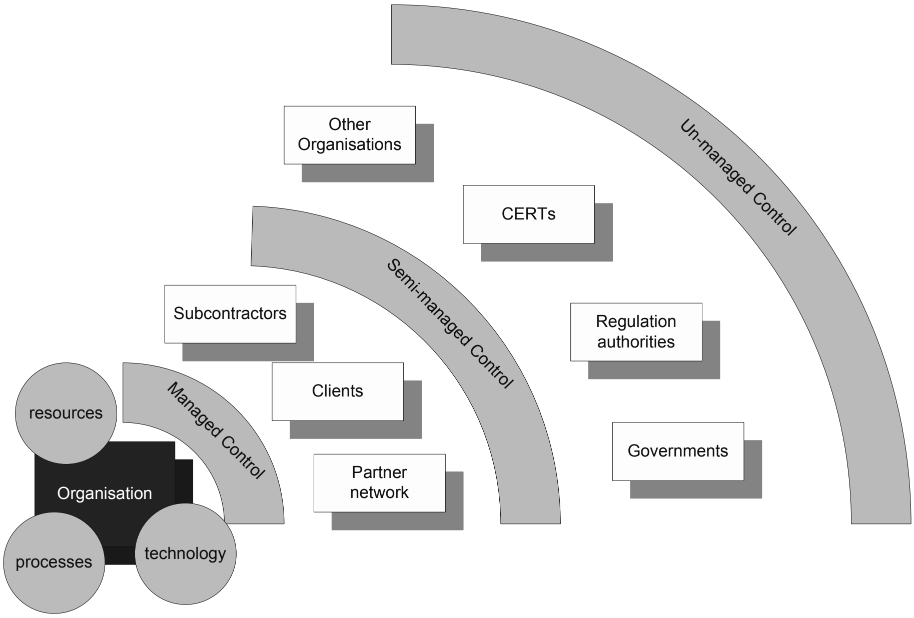


Figure 14. Multi-organizational environment.

Security evaluation, testing and assessment techniques are needed to be able find adequate solutions. Seeking *evidence* of the actual information security level or performance of systems still remains an ambiguous and undeveloped field. To make progress in the field there is a need to focus on the development of better experimental techniques, better security metrics and models with practical predictive power. Security evidence can be used both for quantitative and qualitative analysis methods. Evidence is more useful when they are meaningful for most of the systems lifecycle: during R&D, system implementation and system maintenance phase.

In this research, we use two approaches for obtaining security evidence: *security evaluation* and *security auditing.*

2.4.5.2 Security Evaluation Framework

In the following we discuss a holistic framework [R26], [P11] for model-based information security evaluation or testing of software-intensive systems. Run-time application of results from model-based information security evaluation is presented in [P12].

Role of Threat Analysis

The most important task in the whole process of security evaluation is to identify security risks and threats, taking enough assumptions of the attackers' capabilities into account. A subtask in threat analysis is to identify valuable *assets* that may be subject to security risks. An asset is something in the context of the system that is to be protected. A threat description can be represented, e.g., by *threat* / *asset* combinations. A holistic and cross-disciplinary threat picture of the system controls the development of security solutions. Threats that are possible during the whole life cycle of the system under evaluation must be considered.

It must be noted that the collection of security threats to a system is not static. Security algorithms and other solutions are cracked and new vulnerabilities are found every now and then. Even complete platforms or communication protocol structures can be compromised. As a consequence, a system's threat landscape is constantly changing, possibly reflecting different kinds of trends. A *weak* signal is a factor for change hardly perceptible at present but which will constitute a strong trend in the future. Some weak signals can represent on-going or anticipated changes in the threat landscape. The actual change in time can happen in small steps or in one leap. In the former case, the trend could be exposed, if weak signals presenting the steps could be detected.

Role of Security Requirements

The goal of defining security requirements for a system is to map the results of risk and threat analysis to practical security requirement statements that manage (cancel, mitigate or maintain) the security risks of the system under investigation. Security requirements are constraints on functional requirements intended to reduce vulnerabilities. Security mechanisms are then developed to fulfill the requirements.

The security requirements play a crucial role in the security evaluation. The requirements guide the whole process of security evidence collection. For example, security metrics can be developed based on requirements: If we want to measure security behaviour of an entity in the system, we can compare it with the explicit security requirements, which act as a "measuring rod".

All applicable *dimensions* (or *quality attributes*) of security should be addressed in the security requirements definition.

Modelling Entities and their Cross Relationships

It is obvious that in order to be able to evaluate security systematically, a model of the security behavior of a system is needed. Essentially, the process of security evaluation takes use scenarios and the context of the system into account. In addition to this structural classification of entities, it is important to find the behavioral entities in the system. See [P10] for modeling examples.

Evidence Information

Security evidence is gathered from various sources as input to the decision process of security evaluation. The evidence collection should be arranged in a way that supports evaluation of security behavior and security actions. We classify the types of security evidence information into three categories: measured, reputation and tacit evidence. See [R26] and [P10] for explanations and examples.

Decision Process

The most final phase of security evaluation is the decision process. The overall goal of the decision process is to make an assessment and form conclusions on the information security level or performance of the system under investigation. The decision process can be split into sub-decisions based on the security action model.

The decision process can be carried out in the following way:

1. For each security requirement and security action composition, seek evidence and estimate the probability and impact of that action, taking cross-relationships and trust assumptions into account.
2. Estimate the overall impact of the gathered evidence on each security requirement.
3. Make a decision whether the security of the system with regard to the requirements is at a sufficient level.

2.4.5.3 Security Auditing Framework

To fully cover the risks and especially manage them a framework is needed. The Agile Security Development (ASD) is ISO/IEC 17799-compliant method for developing and continually improving an information security management system. Agility in software development means that the developers only concentrate on functionalities that are needed immediately, delivering them quickly and collecting feedback, and, most importantly, rapidly reacting to business and technology changes [R21]. This definition of agility is the basis of the ASD method. Figure 15 depicts the complete method, which is described in more detail in this chapter.

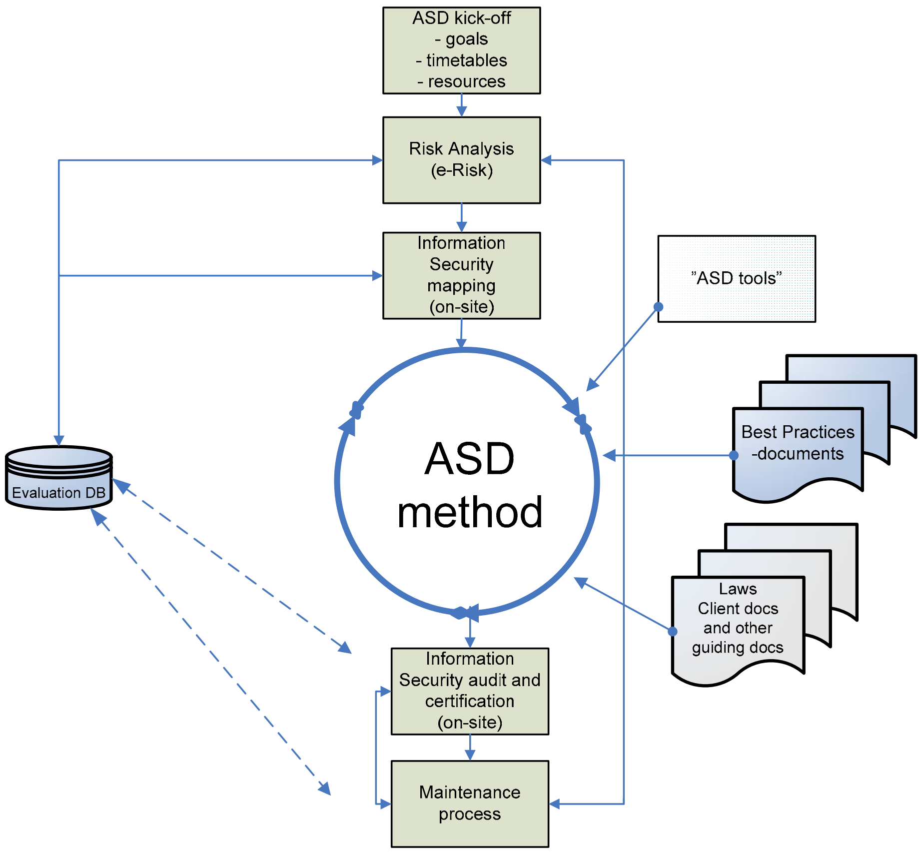


Figure 15. Agile Security Development method.

Walking through the steps of the ASD method starts from the ASD kick-off, where the goals and timetable, as well as the resources for ISMS development, are set. When the scene is set, the actual work starts with the risk analysis step, where the most important parts for development are recognized. Risk analysis is holistic in nature, so it covers both technical and organizational viewpoints [R26]. After that, it is possible to carry out on-site information security mapping to get a clear picture of the current status. The ASD method circle is the core component of the method, where, through multiple iterations, a more complete implementation of the ISMS is achieved. The ASD method may use a repository to store non-conformance reports as well as the status of corrective actions. The stored information can be used for benchmarking. When building an ISMS, the method takes account of legislation, client requirements – such as the client's information security policy – and various kinds of best practice documentation. Figure 16 depicts the components of the ASD iteration cycle. The cycle starts with the management meeting, where the requirements for developing the ISMS are set along with the resources and time schedule. Here, management commitment is ensured and, for example, the measurement requirements are defined. From this point on, based on risk analysis for example, the organization decides where the focus for the rest of the development process shall be and starts the work from the most high-priority parts and even during the work the focus can be shifted to more accurate items. As an example, if risk analysis shows that there are many security incidents related to personnel malpractices, the organization may direct resources to implementing an extensive security education programme and then carry on to the next item found in the risk analysis stage.

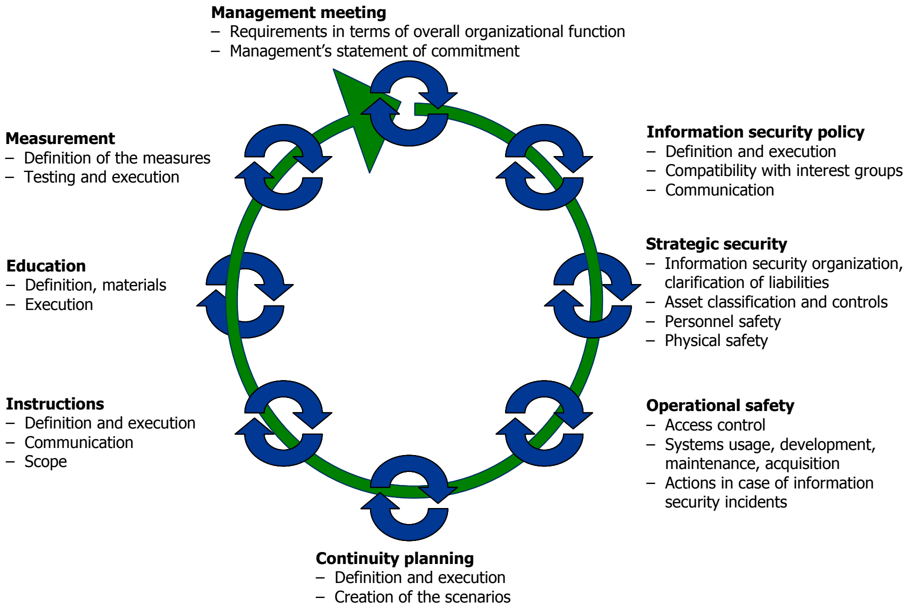


Figure 16. ASD iteration cycle.

The method is fully ISO/IEC17799 compliant, but takes the organization's actual management systems into account, so that ISO/IEC 17799 certification is not necessarily the ultimate target if the organization so chooses. The ASD method supports auditing against the organization's own baseline, which might not be compliant with the existing standards and industry-defined best practices.

After the audit, whether against ISO/IEC 17799 or the organization's own baseline, the ASD method switches to the maintenance process, where regular risk assessments and audits are carried out to ensure continual improvement of ISMS. There, risk analysis is done in each iteration, so that in the next iteration the most important parts are addressed and corrective actions are taken where needed most. This is also in line with ISO/IEC 17799, which requires a coherent and continual risk assessment method to be used as a tool for continual improvement. Agility in the described method is achieved by way of the above-mentioned, but, in addition to this, also by introducing new documentation to the ISMS only where it is necessary and applicable to the organization's own needs.

The ASD method as a whole covers all areas of the ISO/IEC 17799-compliant ISMS from management commitment to business continuity management. The

Agile Security Development method has been piloted in the ITEA SHOPS [R28] project by the Finnish project partners, and is in use in various bilateral projects within Finnish industry.

2.4.5.4 Practical Considerations

In this chapter we shortly analyse the practical implications of security evaluation and security auditing.

Security Evaluation

In the previous section we presented an approach to security modeling and evaluation. Unfortunately, in practice, a thorough modeling of security behavior is only possible in a few ideal cases. Typically, today's software-intensive products are very complex, their functionality is not well documented and often has unknown dependencies. Development of an ambiguous security behavior model at an atomic security action level is a very challenging and time-consuming task.

The practical needs for security evaluation are often limited too. This results to a situation in which we should be able to try to find the security actions that are most critical and most typical. To reach the desired security level it is not important to try and measure every part and component that affects security. Instead, we need enough evidence to make trade-off decisions.

We propose the following process to carry out practical-level security evaluation:

- **Risk and threat analysis.** Carry out risk and threat analysis of the system and its use environment if not carried out before. In real-world engineering, risk and threat analysis are not carried out adequately. Consequently, the set of security requirements might not be sufficient.
- **Define security requirements** in a way that they can be compared with the security actions of the system. Based on the threat analysis, define the security requirements for the system, if not yet defined. These are lacking in many practical systems.

- **Prioritize security and other requirements.** The most critical and most often needed security requirements should be paid the most attention.
- **Model the security behavior.** Based on the prioritized security requirements identify the functionality of the system that forms the security actions and their dependencies in a priority order.
- **Gather evidence** from measured, reputation and tacit security information.
- **Estimate the probabilities and impacts of security actions** based on the evidence.
- **Aggregate the results from the probability and impact estimation** to form a clear picture of whether or not the system fulfils the security requirements and context.

Security Auditing

The auditing itself is not enough in practice since companies often need to enhance their information security management in general. The audits themselves need to be consultative or separate development projects are needed.

2.4.5.5 Security Conclusions

We have discussed the problem of information security evaluation and auditing in the context of software-intensive systems and safety-critical applications. There are no systematic means of carrying out security evaluation. In this paper we have presented a conceptual holistic framework for security modeling and evaluation with some practical considerations. The framework is based on evidence collection and security requirement centered impact analysis.

Information security auditing presents another perspective to gathering security evidence, focusing in management aspects but not forgetting the technology.

This is not a rigorous solution and future work needs to be done on developing a suitable language for expressing security requirements and security behaviour in an unambiguous way. A collection of security patterns would be very helpful in modelling the security behaviour when carrying out security testing or experimentation.

In practical security evaluation, requirements should be prioritized and the system modelled only to the extent needed to conform to the trust assumptions. Full modelling of practical systems is not feasible without automated approaches that are might be very challenging to develop.

2.4.5.6 Security Solutions for Demonstrator

Security solution in the MoNet testbed described in the Figure 2 has several considerations. First, the communication middleware component has a UPnP interface towards the semantic awareness component and a SOAP/HTTP interface towards the content platform. If the SOAP/HTTP interface towards content platform uses public networks, the end points should be authenticated and communication protected with TLS/SSL protocol and public key certificates. Otherwise both UPnP and SOAP/HTTP protocols should only be used in private networks.

Content platform has interaction with communication middleware, broadcasting and DVB-H and HTTP end users. Communication middleware communication uses SOAP/HTTP which should only be used in private networks unless SSL/TLS is used to authenticate end points and protect communication. Content platform communicates internally with DVB-H broadcasting equipment using FTP, which should only be used in private networks, and SSH, which is suitable for public network usage too. No additional authentication or communication protection is needed when broadcasting content to end users with DVB-H and HTTP. If these protocols and interfaces are used for non-broadcast able, private content, then both senders and receivers should be authenticated and communication protected using standard protocols such as HTTP over TLS/SSL (HTTPS) for example. Also PGP like public key encrypted content could be transmitted over plain DVB-H and HTTP channels, but key exchange and distribution is out of this projects scope.

Communication platform will connect end user terminals with the communication middleware and content platforms. UPnP and SIP protocols are used with IPv6, Mobile IPv6 and HIP network protocols towards middleware component. If lower layer protocols do not authenticate end points and protect communication, then UPnP and SIP should only be used in private network. UPnP and SIP should use for example IPSec to authenticate end points and protect communications when communication goes through public networks. Content platform connection uses plain HTTP, which is fine for broadcast content. If private, non-broadcast content is delivered through content platform, then TLS/SSL should be used to protect HTTP traffic and authenticate end points.

The security components aim to protect the users of the system from each other where such protection is necessary. The users shall contain all actors in UML terms, which can be in contact with the system. Traditional design includes only positive user cases and users, but the security work will try to include also some generic negative user cases and their countermeasures to the picture.

As an example, an end-user interface of the system may be exposed to public networks, which also contain hostile users. Hostile users may try to abuse the interface and access the system and even try to impersonate the system to legimite users. Generic countermeasures against hostile third parties in public networks are authentication with tokens, such as usernames and passwords, public key encryption and certificates, and the use of confidentiality, integrity and availability (CIA) protecting access protocols, such as IPSec, TLS and SSH. If these kind of countermeasures can not be used, the system or its sub-systems should be used in private, physically separated networks. For example a home network for private data and applications could use insecure protocols and not authenticate users as long as the network is accessible only within the apartment or building perimeter.

As a general rule, the system SHOULD NOT expose its private information to unauthorized third parties. The system MAY provide broadcast, unidirectional interfaces which deliver content to all possible users. A broadcast interface MAY use insecure protocols which do not authenticate the system to the users as long as the users are aware of this. Users MUST have a way to authenticate the system and the system MUST have a way to authenticate users in non-broadcast interfaces. The users of the system MUST be made aware how their information is handled within the system and which other users and actors may access their information stored within the system. For example if users upload content for broadcasting, they should be aware that the content will be broadcasted and that they will be held responsible if the content is illegal or has other legal consequences.

2.5 Conclusions

The key results of MoNet project are the following.

- Multimedia metamodeling and semantic awareness techniques and solution for environment monitoring
- Content platform with capability for seamless DVB-H – 3G handover during an ongoing multimedia delivery
- MoNet Communication middleware and mobile ad hoc network vertical handover between WLAN and 3G
- Security evaluation and auditing framework
- Integrated Multimedia over Autonomic Network testbed enabling e.g. real-time home surveillance
- 2 journal papers and 10 conference papers
- 2 submitted patent applications.

Publications

[P1] J. Zhou, E. Niemelä, "State of the Art on Metamodel-Driven Multimedia over Mobile Ubiquitous Computing Environment." VTT Technical Research Centre of Finland, 2005.

[P2] J. Zhou, E. Niemelä, "State-of-the Art on Metamodel Driven Multimedia over Mobile Ubiquitous Computing Environments." Presented at Proceedings of the Communication, Internet and Information, CIIT2005, Cambridge, USA, 2005.

[P3] J. Zhou, J.-P. Koivisto, E. Niemelä, "A survey on semantic web-services and a case study." Presented at the 10th international conference on CSCW in design, Southeast University, Nanjing, China, 2006.

[P4] V. Ollikainen, C. Peng. “A Handover Approach to DVB-H Service.” IEEE International Conference on Multimedia & Expo (ICME 2006), July 9–12, 2006, Toronto, Canada. Pp. 629–632.

[P5] V. Ollikainen, C. Peng. “Integration of DVB-T into DVB-H Transmission Systems.” IBC 2006, 7–11 September, Amsterdam.

[P6] D. Pakkala, J. Latvakoski, “Distributed Service Platform for Adaptive Mobile Service.” Journal of Pervasive Computing and Communications. Special issue on Wireless Networks and Pervasive Computing in International Journal of Pervasive Computing and Communications JPCC. 12 p.

[P7] J. Latvakoski, D. Pakkala, P. Paakkonen, “A communication Architecture for Spontaneous Systems.” Wireless Communications, IEEE [see also IEEE Personal Communications], Vol. 11, Issue 3, June 2004. Special Issue on Migration toward 4G Wireless Communications, pp. 36–42.

[P8] J. Holappa, T. Wiander, “Practical Implementation of ISO 17799 Compliant Information Security Management System Using Novel ASD Method.” Technical Report, 2006.

[P9] J. Kajava, R. Savola, “Weak Signals in Information Security Management.” In: Proceedings of the International Conference on Computational Intelligence and Security (CIS) 2005, Part II, Xi’an, China, December 15–19, Springer, 2005. Pp. 508–517.

[P10] R. Savola, “Towards Security Evaluation based on Evidence Collection (Short Paper).” To appear in: Proceedings of the Second International Conference on Natural Computation (ICNC) and the Third International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) 2006, September 24–28, 2006, Xi’an, China, 2006. 4 p.

[P11] R. Savola, “Towards Security Evaluation based on Evidence Collection.” In: Proceedings of the Information Security South Africa (ISSA) 2006 from Insight to Foresight Conference, July 5–7, 2006, Sandton, Gauteng, South Africa, 2006. 10 p. ISBN 1-86854-636-5.

[P12] R. Savola, “Node-Level Information Security Monitoring for Mobile Ad Hoc Networks.” In: Proceedings of the ETRICS Workshop Security in Autonomous Systems (SecAS) 2006, June 6–9, 2006, Freiburg, Germany, 2006. 3 p. ISSN 1613-0073.

[P14] T. Wiander, J. Holappa, “Theoretical Framework of ISO 17799 Compliant Information Security Management System Using Novel ASD Method.” Technical Report, 2006.

References

[R1] DOI, “Overview The Digital Object Identifier System.” 2005.

[R2] B. Smef, “SMEF – Standard Media Exchange Framework.” 2005.

[R3] I. I. Smdl, “SMDL (Standard Music Description Language) overview.” 1995.

[R4] R. Hopper, “P/Meta – Metadata Exchange Scheme.” BBC Technology, Media Data Group, 2002.

[R5] S. Polak, R. Slota, J. Kitowski, J. Otfinowski, “XML-based Tools for Multimedia Course Preparation.” Archiwum Informatyki Teoretycznej i Stosowanej, Vol. 13, 2001, pp. 3–21.

[R6] MRML, “MRML – Multimedia Retrieval Markup Language.” MRML Group, 2003.

[R7] B. Rodriguez, “EDLXML Serialization.” 2002.

[R8] W. C. Smil, “SMIL – Synchronized Multimedia Integration Language.” 2005.

[R9] G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. van der Velde, B. Wielinga, “Knowledge Engineering and Management: The CommonKADS Methodology.” The MIT Press, 1999.

[R10] metamodel.com, “What is metamodeling?” 2002.

[R11] D. B. Lenat, R. V. Guha, “Building large knowledge-based systems: representation and inference in the Cyc project.” Boston, Massachusetts: Addison-Wesley, 1990.

[R12] M. Uschold, M. King, “Towards a methodology for building ontologies.” Presented at IJCAI’95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, Canada, 1995.

[R13] M. Gruninger, M. S. Fox, “Methodology for the design and evaluation of Ontologies.” Presented at IJCAI’95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, Canada, 1995.

[R14] A. Bernaras, I. Laresgoiti, J. Corera, “Building and reusing ontologies for electrical network applications.” Presented at European Conference on Artificial Intelligence (ECAI’96), Budapest, Hungary, 1996.

[R15] M. Fernandez, A. Gomez-Perez, N. Juristo, “METHONTOLOGY: from ontological art towards ontological engineering.” Presented at Spring Symposium on ontological engineering of AAAI, Standford University, California, 1997.

[R16] B. Swartout, R. Patil, K. Knight, T. Russ, “Toward distributed use of large-scale ontologies.” Presented at AAAI’97 Spring Symposium on Ontological Engineering, Standford University, California, 1997.

[R17] S. Staab, H. P. Schnurr, R. Studer, Y. Sure, “Knowledge processes and ontologies.” IEEE Intelligent Systems, Vol. 16, 2001, pp. 26–34.

[R18] A. Gomez-Perez, M. Fernandez-Lopez, O. Corcho, “Ontological engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web.” 1st ed. London: Springer Verlag, 2004.

[R19] W3C-WSA, “Web Services Architecture.” 2004.

[R20] S. O. A. Patil, A. Sheth, K. Verma, “METEOR-S Web service Annotation Framework.” Presented at Proceeding of the World Wide Web Conference, 2004.

[R21] P. Abrahamsson, J. Warsta, M. T. Siponen, J. Ronkainen, “New directions on agile methods: A comparative analysis.” International Conference on Software Engineering, 2003.

[R22] P. Pinder “Preparing Information Security for legal and regulatory compliance (Sarbanes–Oxley and Basel II).” Information Security Technical Report, Vol. 11, Issue 1, 2006, pp. 32–38.

[R23] M. O’Conor, “The implications of Sarbanes-Oxley for non-US IT departments.” Network Security, Vol. 2005, Issue 7, 2005, pp. 17–20.

[R24] OECD, “OECD Guidelines for the Security of Information Systems and Networks.” 2006. [Web document] Available: http://www.oecd.org/dataoecd/16/22/15582260.pdf.

[R25] H. C. Relyea, “Homeland security and information sharing: Federal policy considerations.” Government Information Quarterly, Vol. 21, Issue 4, 2004, pp. 420–438.

[R26] “Risk Management in SMEs.” [On-line] Available: http://www.pk-rh.com/en/index.html.

[R27] Schultz E. E., “Sarbanes-Oxley – a huge boon to information security in the Computers & Security.” Vol. 23, Issue 5, 2004, pp. 353–354.

[R28] SHOPS. “Smart Home Payment Services. Towards the liberalisation of Europe’s utilities industry.” VTT Research Notes 2335. Espoo, VTT, 2006. 35 p. ISBN 951-38-6782-X.

[R29] S. G. Sutton, V. Arnold, “The Sarbanes-Oxley Act and the changing role of the CIO and IT function.” Int. J. Business Information Systems, Vol. 1, Issues ½, 2005.

[R30] U.S: Department of Homeland Security, “The Department of Homeland Security.” USA, 2006. [Web document] http://www.dhs.gov/dhspublic/ [Referenced 24.8.2006].

3. Mobile Active Overlay (MAO)

Arto Juhola, Kimmo Ahola, Mikko Alutoin, Markus Aroharju, Titta Ahola, Sami Lehtonen, Marko Lyijynen, Sami Pönkänen and Pertti Raatikainen

3.1 Background

The growth of the Internet has had a stifling effect on the evolution of its basic technologies. The introduction of mobile Internet terminals, the notions of Quality of Service (QoS) and security are examples of the current development hampered by slow and painstaking standardisation efforts and infrastructure updates. Better ways are being looked for coping with personalised and mobile services of the future.

Active Networking (AN) ideas have been presented to meet the challenge. AN holds a promise of a highly flexible method for experimenting with novel networking ideas and deploying services "on-the-fly". In the area of mobile services, the prospects are particularly promising. MAO (Mobile Active Overlay) is a mechanism to offer a ubiquitous AN service access for mobile users over the present Internet infrastructure. The project MAO was set out to create a practical platform for the exploitation of AN ideas. Lately, the term "programmable network" has been coined to describe the more practical-minded ideas.

3.2 Objectives

The objective of MAO was to develop a platform capable of dynamic, fast and flexible introduction of novel or improved network-oriented services. The platform was to cater especially for mobile Internet users, and to be practicable. The work was based on an idea known as the active or programmable networks, meaning that the behaviour of special programmable nodes of a network can be changed dynamically by reprogramming them ("injecting" code). With this technology it is possible to achieve unprecedented flexibility as far as network services are concerned: Service code can be deployed "on-the-fly" when needed and to those nodes of network where needed.

3.3 Main Results

Of all possible programmable network approaches only those were selected that could be thought to survive and function in an operator environment.

The programmable network node was realised according to forwarding / control plane separation principle as outlined by e.g. IETF FORCES to achieve necessary performance. The security problems were circumvented by allowing only the operator of a network to decide what particular “active” programs may inhabit his nodes, leaving users only the power to request code injection indirectly (by service requests). Incremental deployment of nodes, yet allowing potentially global service access and without touching the existing Internet infrastructure can be achieved using a mechanism with the name of the project (MAO).

The project completed a functional prototype, ready to be utilised in related projects already launched. The prototype consists essentially of an “on-the-fly” configurable router and a programmable server forming together a programmable node.

Impacts

Aside from international conference papers, the implemented programmable node was used in EU IST project CONTEXT, a project that was launched in late 2002 and concerned with automated (mobile) service creation and deployment. Another EU IST project, MAGNET is using the node, now renamed as “LANE” (Lightning Active Network Engine).

Other projects making use of the node are likely to follow, given the practical orientation of the work and the potential for network service implementation. Since the platform is very flexible, the domain of applications is not limited to the so called fourth generation mobile systems (4G); any IP carrying network might be a target. The industrial interest lies mainly on the private companies participating in the said projects.

International and National Networking

There have been intensive contacts to international research institutes and universities from the beginning of the project. Concrete examples of this are that the results of the project are already exploited in EU's Context and Magnet projects. Additionally, we are actively working on a new proposal to continue the research work in this field.

In the national level, we have been discussing with the Finnish industry and network operators to be aware of their interests in the field.

Publications

[P1] M. Lyijynen, T. Koskinen, S. Lehtonen, J. Pesola, "Content Adaptation on LANE Active Network Platform." Proceedings of the 7th International Conference on Telecommunications (ConTEL 2003), Zagreb, Croatia, 11–13 June, 2003.

[P2] P. Raatikainen, O. Martikainen, V. Naumov, "Media Switch for multimedia services." IFIP TC6 Conference on Network Control and Engineering for QoS, Security and Mobility (Net-Con 2002). Paris, France, Kluwer Academic Publishers, Boston, 23–25 Oct. 2002. Pp. 317–328.

[P3] M. Alutoin, P. Raatikainen, "A control interface for router extension." The 21st International IASTED Multi-conference – Applied Informatics (AI2003), Innsbruck (Austria), IASTED, Acta Press, 10–13 Feb. 2003. Pp. 697–702.

[P4] M. Palola, M. Majanen, "Performance of distributed calculation in an active network." The International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'02), Las Vegas, Nevada, US, CSREA Press, 24–27 June, 2002. Pp. 1056–1062.

[P5] S. Lehtonen, K. Ahola, T. Koskinen, M. Lyijynen, J. Pesola, "Roaming active filtering firewall." Smart Objects Conference (Symposium Objects Communicants, sOc'2003), Grenoble, France, 15–17 May, 2003.

4. Best Effort Traffic in Optical Packet Switched Networks (Betraffic)

Pertti Raatikainen, Mikko Alutoin, Kristiina Hytönen, Riikka Lemminkäinen, Petteri Mannersalo, Ilkka Norros and Kari Seppänen

4.1 Background

For the time being, Internet Protocol (IP) is seen as the unifying network layer solution, which can carry almost any service over any link layer. However, IP is lacking features that are essential for successful delivery of services that require constant bit rate, low delay and low delay variation. Basically, the IP networks offer "best effort" type of transport service, which roughly implies that packets end up to their intended destination if everything goes well. If no extra measures are taken, end-to-end delays can vary and no bit-rate guarantees can be given to any service.

Another major trend in the communications networks has been the upswing of the optical transport. Although optical fibres are already common transport media and optics is anticipated to become the prevailing technology in the wire-line network, networking concepts that take optical media inherently into consideration are in the developing phase. For example, IP layer is not able to make full use of the huge transport capacity of the optical fibres. This can be seen, e.g. in unreasonably coarse granularity of transport resources and mismatch in IP packet and optical signal routing.

So, there is an obviously need to develop interworking mechanisms between the optical and IP layer. These include, e.g., routing and wavelength assignment, resource allocation and connection control, network management as well as security of communication. Since there is always some adapting layer between the IP and optical layer, the interworking naturally involves some third layer, which complicates the matter. The interworking solution should consider needs of the IP and optical layer as well as those of the layer residing between the two. For example, some adapting layers, such as SDH/SONET, implement efficient fault location and restoration mechanisms that are totally lacking from IP.

This leads to a new problem: reliability of communications. Transport concepts usually deployed in the optical fibres implement mechanisms to cope with fault situations. If the IP layer is to carry all sorts of traffic and utilise the high bit-rates of optical fires, the IP layer should also include some reliability measures. For example, disturbances on the IP layer reflect to the adjacent protocol layers (above and below the IP layer) thus possibly triggering network management operations on several layers. Some of the triggered operations consider status and alarm information, received from other management layers, and synchronisation of the transfer of this information is of crucial importance. If the synchronisation is lost, an avalanche of control messages is generated leading to conflicting management operations and ultimately to an uncontrolled state of the network.

Despite the technology and physical availability of the communications equipment, reliability of the Internet is also affected by human intervention. Well known human caused disorder include, e.g. worms, viruses and malicious attacks to central network nodes. So, there is an obvious need to identify the most vulnerable points of the Internet and find mechanisms that may cause major problems.

4.2 Objectives

The first objective was to define the optical network layer, residing between the IP and optical layer, and draw up its most obvious functions (highlighting the unique tasks it has to carry out in adapting IP to the optical media). The second objective was to survey and develop mechanisms to enable proper interworking between the IP and optical layer. The interworking was to include functions such as transfer of information about fault situations and recovery actions.

Since the reliability (and especially availability) of an optical IP network turned out to be a problem, the last objective was to identify the most vulnerable points of the Internet and develop reliability models for the optical IP network. The aim was to concentrate, especially, on the protection and restoration techniques, deployed in the existing transport networks, and their impact on the IP layer reliability. The first step was to cover reliability models for limited size networks and the final goal was to develop a reliability model for a large IP network.

4.3 Main Results

The main results of the project are summarised in the list below.

- Definition of an optical transport layer, which covers functions of the adaptation layer residing between the optical and IP layer.
- Survey and development of methods to intensify interworking performance of the IP and all-optical transport layer.
- Definitions for all-optical IP network reliability.
- Development of a layered reliability model for an all-optical IP network.

Impacts

The developed optical transport layer and IP-to-optical interworking methods have been notified by some of VTT's industrial customers and a nationally funded follow-up project has been launched. Another project that studies transport capacity provision methods in optical access networks has also been started. The developed reliability model of an all-IP network inspired a multiparty follow-up project, which concentrates on analysing reliability of the future packet based mobile networks.

The results of the project have been disseminated in various seminars and workshops. Based on received feedback, we anticipate that the developed know-how will generate many other possibilities in the field of optical communications and in the field of networking reliability.

International and National Networking

During the course of the project we have constantly exchanged information with the Finnish industry and network operators. The aim has been to inform possible customers about the ongoing work and obtain some feedback. There have also been scientific contacts with Finnish Universities, e.g. Helsinki University of Technology (TKK), University of Helsinki and University of Jyväskylä.

We have also had contacts with some European research institutes and universities in finding out co-operation possibilities. Some concrete project proposals are underway and there is already scientific co-operation with some universities. Additionally, our research team is a partner in EU's Euro-NGI network, which also will study the issues touched in this project.

Publications

[P1] K. Kaario, P. Raatikainen, M. Pääkkönen, T. Hämäläinen, "A Lightpath Allocation Scheme for WDM Networks with QoS." Proceedings of the 8th IEEE Symposium on Computers and Communications (ISCC'2003), Kemer-Antalya, Turkey, 30 June – 3 July, 2003. Pp. 684–689.

[P2] P. Mannersalo, P. Raatikainen, "IP over Fibre – the optical network layer." Research Report, VTT Information Technology, January 2003. 21 p.

[P3] P. Raatikainen, "Best Effort Traffic in Optical Packet Switched Networks." Research Report, VTT Information Technology, December 2002. 20 p.

[P4] P. Raatikainen, "Dimensioning of Switch Fabric Interfaces of High-Throughput Routers." Proceedings of the IASTED International Conference on Communications and Computer Networks (CCN 2002), Cambridge, MA, USA, November 2002. Pp. 269–274.

[P5] K. Kaario, T. Hämäläinen, P. Raatikainen, "Adaptive Parameter Setting for QoS Aware Load Balancing Algorithm." WSEAS Transactions on Communications, Vol. 1, Issue 1, 2002, pp. 144–149.

[P6] P. Raatikainen, H. Pentikäinen, K. Ahola, T. Frantti, A. Juhola, A. Mämmelä, J. Lähteenmäki, S. Pönkänen, A. Rova, T. Sukuvaara, "Interoperability and Mobility in Future Networks." In: Communications Technologies – The VTT Roadmaps. Ed. by M. Sipilä. VTT Reasearch Notes 2146, Espoo, 2002. Pp. 9–25. http://virtual.vtt.fi/inf/pdf/tiedotteet/2002/T2146.pdf.

[P7] P. Raatikainen, M. Wikström, T. Hämäläinen, “Response Delay of a Fair QoS based Caching Method.” Proceedings of ICT2002, Beijing, China, June 2002. Pp. 1056–1060.

[P8] P. Raatikainen, M. Wikström, T. Hämäläinen, “Performance Simulations of a QoS Aware Caching Method.” Second International IFIP-TC6 Networking Conference (Networking 2002), Pisa, Italy, May 2002. Pp. 1202–1207.

[P9] P. Raatikainen, M. Wikström, T. Hämäläinen, “A QoS Based Caching Scheme and Simulations.” Proceedings of the IASTED International Conference, Applied Informatics (AI2002), IASTED, ACTA Press, Innsbruck, Austria, February 2002. Pp. 289–294.

5. Design of All-IP Mobile Network (DAIMON)

Pertti Raatikainen, Wenche Backman, Jari Hautio, Annemarie Hjelt, Tero Haatanen, Seppo Horsmanheimo, Paul Kemppi, Jorma Kilpi, Krzysztof Kordybach, Jari Laarni, Yanli Li, Eija Myötyri, Ilkka Norros, Leena Norros, Unto Pulkkinen and Tapio Suihko

5.1 Objectives

The telecommunications infrastructure is undergoing two fundamental changes: the circuit switching technology is giving way to the packet switching one, and the user access is becoming wireless, even for broadband. The present mobile networks belong to the category of circuit switched networks, but extensions have been made to allow packet based communications as well, e.g. GSM/GPRS and UMTS. Circuit switching can in principle be replaced by the Internet Protocol (IP) and related communication protocols that enable to carry all kinds of services over most varying networks. On the other hand, the first purely packet based wireless networks, the WLANs, still lack efficient mobility features. Smooth and seamless handovers from one wireless access network to another type of wireless network, so called vertical handovers, provide challenges for ongoing research projects.

Due to the huge investments tied to the old established infrastructure and the fact that the telecom infrastructure is worldwide, the change does not take place overnight. As the technology evolves, so does the network architecture: new network element and functionality appear affecting the way the network elements communicate and connect to each other. When all sorts of networks, including the future sensor and personal area networks, really start to interwork, we can finally speak about a ubiquitous network environment. At some phase of development these networks will most probably allow people to be always connected to the communications networks and via the best available connections.

In this vision, the ubiquitous network environment provides all-time access to various resources everywhere for everybody. It also offers people new ways of coordinating everyday life activities. It is often envisioned that the future users would not need any perception of the network environment, but in fact it is far from clear whether this goal of complete transparency of the network is achievable or reasonable.

The future possibilities of network usage may also include negative impacts that need serious consideration. For example, people's life may become more and more fragmented and less predictable due to increasing interruptions and changes of attention. Moreover, there is a threat of collapse of social boundaries, if people's private life becomes more public, causing identity problems and stress. Because of these challenges, the implications of wireless technologies for the individual and society should be carefully considered.

The main objective of the DAIMON project (Design of All-IP Mobile Networks) was to

- analyse and develop present and future architectures of the infrastructure connecting the base stations of various wireless networks to the Internet, and to make focused studies on various relevant aspects of these architectures.

Another objective of the project was to

- create new kind of interdisciplinary collaboration around the topic by the inclusion of expert viewpoints of systems dependability and human-technology interaction.

5.2 Approach and Methods

In the DAIMON project, we saw that the challenges we meet in the development of the future wireless network architecture arise from several kinds of heterogeneities:

- heterogeneity of technologies: interworking, vertical handovers
- heterogeneity of geography: it is not possible to provide "unlimited bandwidth everywhere" – how, then, should the networks be designed

on peripheral areas in order to lower the significance of geographical location instead of increasing it even further?

- heterogeneity of the users: at present and in future, people's readiness to use new communication tools varies hugely
- heterogeneity of network operators: how can the user see the network environment as a unified whole, although it consists of networks operated by many independent companies?

To tackle such a broad spectrum of problems, a multidisciplinary approach was considered necessary. We aimed at creating mutually fruitful interaction of experts representing different disciplines, e.g., integrating a serious analysis of the user point of view from an early phase of the work. The implications of wireless technologies for the individual and society should be carefully considered. Wireless network services should be analysed in terms of their ability to support the co-ordination of everyday life, and since consumers perceive and evaluate mobile services in the context of everyday life activities, service quality should also be assessed in mobile contexts of use. To that end, a broader systemic perspective was taken in the DAIMON project, and the development of network architecture was also studied from the perspective of dependability and end-users' experience.

Thus, the DAIMON project studied the architecture of network structures providing technology-independent wireless access to packet-based multimedia services, together with several specific aspects of such an infrastructure. In practice, the work focused on the following topics:

- survey of existing implementations of large-scale wireless packet access
- radio resource allocation mechanisms
- conceptual frameworks for analysing the human user's relation to the network
- characterization of wireless network traffic
- network-based positioning techniques
- alternative architectures for connecting heterogeneous access networks to the Internet
- dependability studies of these architectures.

The approaches and methods used in these specific studies are described below.

Existing implementations: The study was made as a literature review, with emphasis put on the network structure chosen for wide area WLAN implementations and other technical details. The material was mostly found on the Internet but results from interviews with network planners and users were also taken into account.

Resource allocation mechanisms: The study of UMTS resource management was based on literature as well as on simulations performed with our own tool. No trial network was available. In case of purely all-IP networks (WLAN and WiMAX) the research was based on a literature review since neither a simulator nor a trial network was available. Given the long background in resource management in UMTS networks at VTT, it was possible to identify problems and trends in resource management also in all-IP networks.

Traffic characterization: Using a large high-resolution sample of a commercial operator's GPRS traffic as empirical data, advanced statistical methods were applied and developed further.

Positioning techniques: A location algorithm for indoor environment was developed and evaluated. An existing terminal measurement software tool was enhanced and a new test bed for ubiquitous terminal assisted positioning was designed and implemented. In addition, trials (WLAN and RFID) were performed and analysed. A literature survey of indoor positioning techniques was conducted.

Architectures: The study of the future network architecture adopted the ITU-T's NGN and 3GPP's network concepts as the starting point. These define a large number of different network elements necessary in the united circuit-packet switched networks. The target was to see beyond the circuit-packet network era and foresee which of the existing network functions will survive and which functionality should be added to the future networks. The developed ideas were compared with ideas found in international publications.

The architecture of the mobile backhaul network was selected to be studied more thoroughly. It was assumed that if the mobile access data rates increase significantly, all sorts of data (also other than mobile data) are carried by the backhaul network and the different users and categories of data should be identified and separated by the backhaul transport techniques.

Dependability studies: Traditional structural reliability analysis was applied to the schematic architectures considered. The broader view on dependability aspects of IP networks, originated in the BETRAFFIC project, was integrated into the work on human-technology interaction.

Human-technology interaction: In this case, a somewhat more substantial methodological discussion is motivated. User-oriented design has already for some time been an accepted goal for the design of human-computer interaction and user interfaces for everyday appliances. Another user-centred approach is the so-called human factors perspective which plays an important role in the design of human-system interfaces in high-reliability environments. Developing a user-centred perspective to the design of large infrastructures such as IP networks could make use of the experience gained in both above mentioned user-centred design traditions. Actual attempts to realise this possibility in the design of IP infrastructures are currently very few. In the DAIMON project we made such an attempt. We saw that it would be beneficial already in the conceptual phase of the design of all-IP infrastructures to try to foresee what the usage of the network would require for the technology and its implementation. Hence, an appropriate approach also in the all-IP network development should give advice *how* to define requirements in the development of the network for different contexts and services, *not* to try to define *what* features the solutions should specifically have.

Outlining the user perspective was seen to be mainly a theory and concept building task. The goal was, in collaboration with the experts of network and telecommunication technology, to conceptualise the intrinsic features of all-IP networks. Our attempt was to outline a new approach to comprehending user experience and acceptance. We accomplished literature studies in the areas of usability and user experience of mobile ICT and further studies in media theory. We also took advantage of the possibility to integrate the project's measurement-based studies on the round-trip time variability into the study. The findings in the literature and own measurements were to be synthesised into a model of user acceptance and into further hypotheses concerning the formation of user acceptance.

5.3 Main Results

We present the results of the DAIMON project in two parts, called respectively "User perspective" and "Network perspective". The first part proposes a model of user experience incorporating an integrated view on the technical vs. perceived Quality of Service, and a broader media-theoretic conceptual framework. The second part presents the results concerning network technologies, architectures and traffic.

5.3.1 User Perspective

5.3.1.1 Social Implications of Wireless Technology

Castells et al. have proposed in [R6] that new forms of space and time emerge due to wireless technology. According to them, wireless communication both 'homogenizes' space and creates a new practice of time. The space of social interaction has to be defined differently when people that are located in different places communicate with each other and use their present place to build a network of communication with other places. In this context service quality is appreciated in terms of support for the building of communication networks. Services should, for example, provide support for rendezvousing, "the informal, geographical co-ordination of small groups of people" ([R6], p. 233).

Also new forms of time are emerging. There is a de-sequencing of social action based on time compression and by random ordering of the moments of the sequence and a blurring of time due to de-sequencing of activities that is allowed by space of persistent connectivity ([R6], [R11], [R24]). Mobile communication makes possible to fill time with social and other practices, i.e., it makes possible to insert communication and use of information services in those moments when other activities cannot be conducted. Therefore, service quality should also be appreciated on the basis of supporting this type of use of mobile technology.

All in all, wireless network services are perceived and evaluated in terms of their ability to support the development of communication networks and the co-ordination of everyday life. Since people always perceive and evaluate mobile services in the context of everyday life activities, service quality and acceptability should also always be assessed in the context of everyday living and in mobile contexts of use.

5.3.1.2 Quality of Service

Quality of service (QoS) is a set of judgments of how good the service is. It has been defined as "the collective effect of service performance which determine the degree of satisfaction of a user of the service" [R14], and "the degree of conformance of the service delivered to a user by a provider in accordance with an agreement between them" [R14]. In [R15] the definition of quality of service is based on three criteria, 1) *speed* indicates aspects of temporal efficiency of a function, defined on measurements that are made on particular time intervals, 2) *accuracy* indicates the degree of correctness, based on ratio or rate of incorrect realizations of a function (losses), and 3) *reliability* indicates the degree of certainty with which a function is realized, which is related to dependability of the system.

Typically, a distinction has been made between subjective and objective QoS in which perceived and assessed QoS are included in subjective QoS [R12]. According to Hardy ([R12], p. 5) objective or intrinsic quality of service is related to technical design of the communication network and terminations and provisioning of network accesses, terminations and switch-to-switch links. Perceived quality of service results when the service is actually used, and it determines whether people will be satisfied with the service.

Objective Quality of Service

For streaming applications, audio and video, a user can always hear or see the bad quality. There exist natural variables like, for example, delay or delay variation and well known bounds for these variables such that if these bounds are not exceeded then a user does not recognize any problems with the streaming application.

The knowledge and understanding of the factors that currently affect to experienced service is a natural cornerstone when trying to say something about the QoS of future users. Still today, TCP is the dominating transport protocol and this situation is likely to continue in the near future. For a feedback based protocol, like TCP, *the round-trip time (RTT)* affects to the data transfer rate that the TCP offers to the application layer(s). TCP is also used for streaming applications since its simplified version, called User Datagram Protocol (UDP), has no flow control (feedback) features and it is not intelligible to re-program any flow control features at the application layer.

<u>RTT variability</u>

For a feedback based protocol, like TCP, the *RTT variability* is one of the main factors which affect to the available bandwidth offered for the application layer protocol. Thus it is also a thing that the end user experiences about the service. Roughly expressed if RTT stays well below 1 second during a data transfer (flow) then a user may experience a continuous though perhaps slow transfer rate but if RTT gets sometimes, even only once, values over 1 second then a user typically notices a break in the data transfer. A thorough measurement based study of the RTT variability has been finalized within the DAIMON project [P7]. The data used in the analysis was obtained from the Finnish mobile operator Elisa in an earlier, TEKES funded project PAN-NET. The results of [P7] can be summarized as follows.

In current mobile networks, the RTT variability is mostly due to a deterministic reason, namely the slow access rate and buffering of packets before the slow radio access link. The load by other users of the same radio interface and the distance to the base station may be reasons but not very dominating or typical ones. Actually, the other main reason for RTT variability is the user itself. Namely, with broadband access an experienced user may be accustomed to use several applications simultaneously but the same behaviour does not fit well to the combination of slow access speed and relatively low (possibly even battery limited) CPU capacity. With such a combination simultaneous data transfers of different applications disturb each other significantly. Especially simultaneous up- and downloading caused severe problems for TCP. The methodology in [P7] was based on applying signal processing methods, Fourier and wavelet analysis to the measured RTTs of TCP flows.

The effect of the bandwidth that the application really gets

Dependence between downloaded file size and access rate has been studied a few years ago in Würzburg, Germany, and also in VTT in the cases where reliable knowledge of true access rates were directly available. See [R17] and [R32] for more details. These older results on wired (dial-up) networks confirmed the quite obvious hypothesis that the higher the access rate is, the higher the probability of downloading larger and larger files is. This is a probabilistic, not deterministic statement and it must be emphasized that very large downloadings occur also with very slow access rates. This latter is just more seldom but still a regular phenomenon and is statistically reflected as heavy-tailed character of the file size distribution.

In wireless connections the best access rate and the true data transfer rate that the application really gets may differ quite significantly. More dedicated studies of the dependence between a web content (file) size and the (mean) rate that the TCP protocol offers for downloading of the web content have been initiated in the DAIMON project ([P6], [P15]). One preliminary motivation for this research issue was to understand how much the real transfer rate that the application gets affects to the user behaviour. The preliminary results of [P2] were based on the approach of expressing the bivariate distribution of a TCP flow size and the (mean) obtained rate in terms of a so called copula. The work started in [P6] is still in progress, aiming now more to the analysis of the particular effects of TCP and, actually, excluding the user behaviour. In [P15] a different, purely statistical approach to the same problem has been taken which includes the user behaviour. The work of both [P6] and [P15] continues in the TEKES funded project ABI.

Subjective or Perceived Quality of Service

Perceived quality of service is defined as a subjective quality of the service perceived by the user. It includes dimensions such as loading latency, sound and image quality and perceived reliability. There is always a trade-off between complexity (how many network parameters are taken into account) and usefulness (capturing accurately the interaction between a mobile network and perceived QoS [R16]). Overall Perceived Quality of Service (OPQoS) can be thought as a weighted sum of individual quality factors.

Perceived QoS is typically assessed by such measures as Mean Opinion Score (MOS) and Task Oriented Performance Measures (TPM). When using MOS, users express their judgments according to a particular scale; TPM present users different levels of stimuli and the outcomes are measured objectively. Table 1 presents some measures of perceived QoS classified according to whether they are subjective or objective and qualitative or quantitative. We aim at developing a dialogic method for investigating people's attitudes and experiences of information technology. The method is based on such techniques as focus group –methodology, repertory grid analysis and dialogic intervention methodology.

Table 1. Possible measures of perceived/experienced QoS and appraisal of service.

	Subjective	Objective
Quantitative	Post-test rating scales Continuous assessment techniques – e.g., SSCQE Psychophysical methods – e.g., MOS	Performance measures – e.g., TPM Behavioral measures Psychophysiological measures
Qualitative	Continuous verbal reporting Ethnographic methods Focus groups Etc.	Course of action analysis

5.3.1.3 Extended Model of Quality of Service

A basic problem with all previous models of perceived quality of service is that they do not consider to a sufficient degree the role of context and usage situation. Perceived quality of service is critically dependent on by which way the service functions in a particular context. We have to 'contextualize' QoS models and systematically analyze how people use mobile technologies in different usage situations. For example, we have to analyze which kind of tasks people have and ways they accomplish these tasks without/with mobile devices. User acceptance and quality of service have also typically addressed from the perspective of individual users. However, since the use of wireless technology is typically a collective enterprise between groups of people, we should pay more attention to the needs of communities of users. To that end, we need a systemic analyses and studies of how people and communities they live in experience mobile technologies in different contexts.

Based on a review of literature of perceived QoS an extended model of QoS (exQoS) has been outlined (see Figure 1).

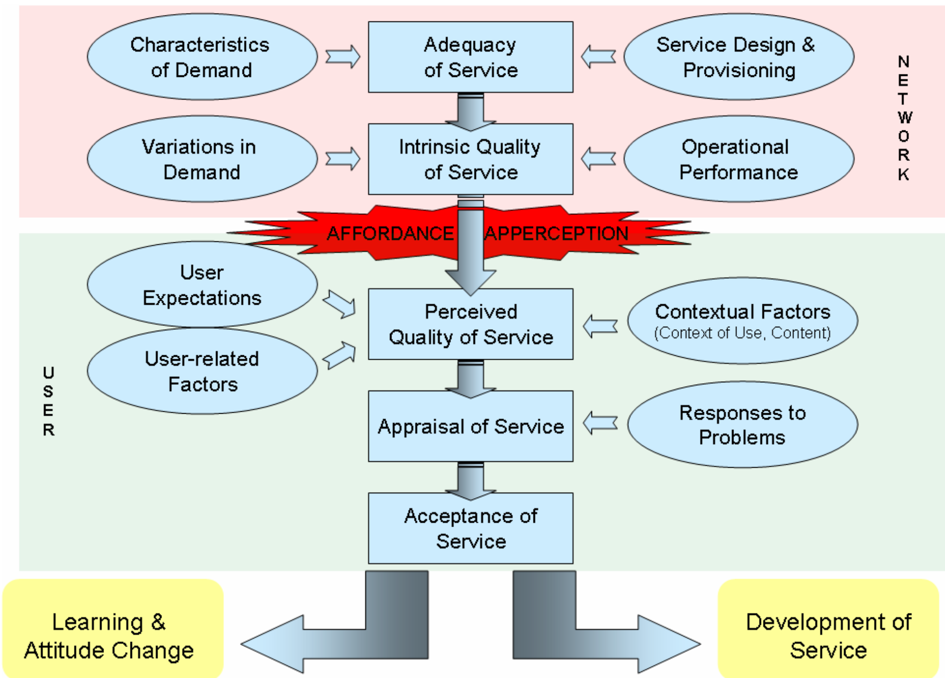


Figure 1. An extended model of quality of service (exQoS). The model is modified from [R12].

According to the model, appraisal of service (AS) and acceptance of service result when the user makes an evaluation on whether the quality of service is good enough to rely on. Te appraisal of service is shaped by users' experiences while interacting with the service. It concerns the question of what is the meaning of the service to the user. Appraisal is influenced by user experiences in a particular context and in a particular activity. Several authors have presented lists of features that are characteristic to mobile contexts (e.g., [R20], [R24], [R29]). Mobile devices are used both in work contexts and during leisure activities. Major characteristics of mobile contexts are that there are typically frequent interruptions, temporal tensions, and frequent side courses and sidesteppings. People also typically have to manage multiple tasks simultaneously.

Typically, there is no linear relationship between intrinsic (objective) quality and perceived (subjective) quality. Neither can users distinguish between network and application/terminal problems. For example, Koivisto and Urbaczewski found that users could not separate network performance into multiple factors like connection establishment time, bandwidth and release time but they were considered as a single factor [R18]. Additionally, the relationship was shown to be application specific. It is highly probable that services for new kind of smart, pervasive environments will cause new problems for the measurement of perceived QoS.

In order to better comprehend the connection between the engineering quality features of the network and the experience people have when they use the services build on the network we decided to draw on the idea of the psychologist James Gibson [R10]. According to him the human actor's perception of the environment is not a subjective representation of the physical parameters that are used in a scientific characterisation of the environment. Instead people perceive the environment in the form of holistic attributes that Gibson called affordances [R10]. These are latent action possibilities in the environment, independent of the individual's ability to identify these possibilities.

The idea is to identify what are the generic attributes of the mobile all IP-network that provide new types of possibilities for the user. In the interdisciplinary discussions and work with the DAIMON projected we have arrived at three generic attributes of mobile network technologies which we assume to have implications on service quality and acceptance. These are connectivity, contextual awareness and adaptability:

> ***Connectivity*** ("being connected") means that the user is Always Best Connected (ABC-concept) to the service, that is, the user has an access to the mobile network through the best possible wireless interface. Users are thus always connected to services in the best possible way, and the network offers flexible usage of services anytime and anywhere (mobility).

Location awareness ("being located") means that a mobile device can be sensitive to its current location and should function accordingly. Location awareness is one form of contextual awareness, the application's ability to react in a sophisticated way to situational information. Basic forms of contextual awareness are location awareness, environmental awareness, mobility awareness, and user awareness.

Adaptability ("being served") and especially self-adaptability, refers to "to any variants tending to modify the end-to-end communication, preventing the user from feeling changes" ([R3], p. 2051)

Connectivity affords a possibility to being connected, that is, experience or belief of being connected to other people, locations, services etc. anytime and anywhere. Location awareness provides a possibility to being located, that is, an experience/belief of being located by other people/services where ever the user is physically located at any given time. Proactiveness and adaptability, in turn, provides prerequisites for being served, which includes the experience or belief of being smoothly and continuously served and supported.

Concerning the appraisal of service, and finally the acceptance of service, the critical questions are: 1) how generic affordances (to be connected/located/served) influence appraisals; 2) what is the role of perceived quality in appraisals; and 3) how users' experiences shape their appraisals.

In Figure 2 some aspects of user experience that may be relevant to user acceptance are shown. An interesting question is which type of effects the generic attributes of smart/mobile computing environments may have on these attributes. Considering connectivity and mobility, inter-technology roaming and limitations of portable devices may cause transient fluctuations in QoS, and this, in turn, would cause reduced feeling of control/trust and feelings of irritation in users. Contextual awareness by making proactive/predictive services possible may obscure limitations in intrinsic quality so that limitations in QoS remain unnoticed. On the other hand, services may be privacy-intruding and reduce feeling of control/trust. Connection speed is also critical: Services must be delivered just on time, otherwise they may elicit feelings of irritation.

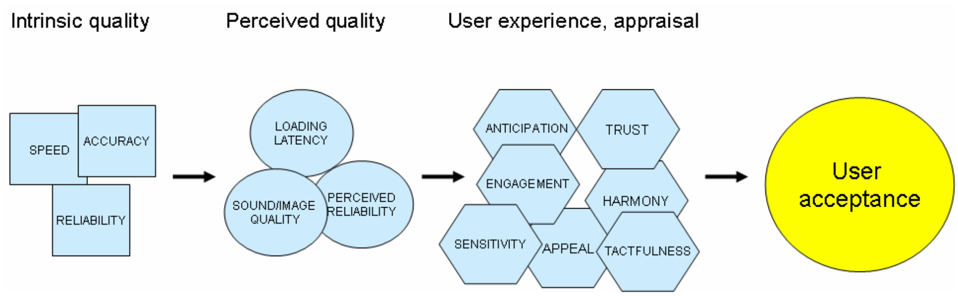


Figure 2. Examples of possible dimensions of user experience that may be related to intrinsic/perceived QoS.

5.3.1.4 Apperception of IP Connectivity – Creating Network Literacy

With reference to the "Extended Model Quality of Service" (exQoS; Figure 1) we raised questions concerning features that would have an influence on the user appraisal of service. Among other things we asked how the generic affordances of mobile technologies, connectivity, contextual awareness and adaptability, would influence appraisals. In this section we shall return to this question.

We shall discuss all-IP mobile network as a resource for different wireless technologies for diverse everyday usage. From this resource perspective the affordance of *connectivity* is the central generic attribute that we have to consider. The two other affordances, i.e. location awareness and adaptability, relate closer to services that are build upon IP network. Hence, they make use of connectivity or add further features to it.

The network environment affords the feature of connectivity that can provide possibilities for the users to act and develop their activities. According to Gibsonian theory, it further holds that in order for the possibilities to exist, the users need to be in the position to apprehend i.e. grasp these features. We consider this ability to "read" the environment as *network literacy*.

In this section we shall elaborate the connectivity affordance of the IP network, and then propose some basic features of what could be considered as network literacy and how to acquire it. The section will end with some notes for future work in the area.

Mobile Network as a Tool and Medium

Taking a user perspective denotes analysing the use of technology in different life activities in which peoples' practices and cultures are formed. Technology is one part of these practices and culture. It is possible to approach technology at least from two different perspectives, i.e. technology as a tool for human activity and technology as a medium that enables communication and shared consciousness. These frames denote different theoretical traditions of social science research. Both are relevant for considering the role of ICT technology in general and, as we see, IP network in particular.

When considering IP network as a *tool* we take a system-oriented point of view to human activity. According to the cultural-historical theory of activity ([R19], [R31]) human conduct could be considered as a system of mediations between the subject, object and the human community. Tools – in this case the network – have three basic functions. These are the instrumental, psychological and probably communicative function. The instrumental function is the most evident one relating to a tool's role to shape the environment in a targeted way. The second function, the psychological function is less evident but quite important. It refers to the fact that by using and developing tools, including concepts, human beings create an external factor that influences their own mental structure and bodily behaviour. In order to a hammer to be a tool there must be the instrument but also the ability to use the instrument [R26]. However, the hammer is a full-fledged tool due to being developed as such in a community, members of which all know how to use the tool, what it means to use the tool, and what is good usage of the tool etc. The tools crystallise practices of usage and mediate them to further people. In this sense tools clearly have a third, the communicative and a sense making function.

The communicative function of tools and the interaction of the functions is even emphasised when the tools become reliant on the ICT technology that basis on electricity and digital technology. ICT enables efficient and intelligent control and monitoring of systems, mediates information effectively and provides a universal way of interacting with the environment that is shared by nearly everybody in every domain of the society. Due to this ICT has acquired a consciousness shaping role [R27]. Hence, it becomes evident that technology could and should today be also dealt with as a *medium*. Hereby we enter to the domain of media theory.

In media theory it is less usual to consider the role of technology in mass communication, and to see that all technologies would have the communicative and consciousness shaping function typically attributed to mass media ([R21], [R25]). One of the media theoreticians who deliberately dealt with a broad concept of media technology and the role of media in shaping human perception, activity, consciousness and culture was Marshall McLuhan (1911–1980). McLuhan focused on the connection between technology and human perception. He was particularly known of the idea of technology as an extension of man, and electrical media that of the nervous system. He identified the background role of media in shaping the way people comprehend their environment. This idea was expressed clearly in his slogan "Medium is the message" or "Medium is the massage". both indicating that media shape humans beyond the content that is mediated.

McLuhan was criticized by colleagues of technological determinism and was almost forgotten after the 1970's [R21]. Today McLuhan has experienced a renaissance when IP network has verified and made real his vision of the electrical network that is implemented everywhere and used in numerous services using multimodal interfaces. While aware of the critique to some aspects of McLuhan's thinking we appreciate the insightfulness of his ideas. We are open to make use of these ideas when we attempt to conceive the role of IP in the activities of people. It is not possible to explain in detail McLuhan's abundant ideas in this connection. We restrict ourselves to demonstrate his main idea of how to identify the performance shaping nature of new technologies.

Connectivity is the Central Affordance of IP Medium

Electricity was the medium extensively considered by McLuhan. This medium developed first in the form of telex, then telephone, radio and television. Television was the last electric tool McLuhan himself experienced. In this medium he already saw the basic intrinsic feature of electrical media to connect with the human nervous system and to re-establish the synaesthesia or integration of senses that disappeared with the invention of phoneme based alphabets, linear writing and dominance of the visual modality. McLuhan and McLuhan see the electric media to resemble the tactile sense. They see this sense to have an integrative role in perception. The idea that McLuhan was advocating was that, due to the physical features of producing the TV image and the resulting low-definition image, watching TV is closer to tactile than linear visual sensing. This would be even strengthened by the use of colours that facilitate a tactile way of interacting with the object.

What has the computer technology and its connection to the IP added to this picture? First of all, an increasing part of the objects that people handle with are transforming to immaterial digital entities that can be called files. This concept was first used in restricted sense by computer professionals but today people are accustomed to think music performances and films as files. This is just an example of the forthcoming change in the ontology of the human world and environment. Connecting the internet to PC, or to other user appliances, and the availability of multimodal forms of representation appear to provide a way of ripening the possibilities of electricity, in analogy to extension of the role of alphabet medium via the print house technology. It should be emphasised, however, that the transformations of ways of perceiving the environment and that of practices is a very slow process. It has been shown e.g. that the transition from a predominant speech-mediated communication to alphabet-mediated written communication took hundreds of years [R28].

The medium nature of IP network is supported by the fact that IP is a generic and independent technology that supports innumerable services. IP network is distinct from any earlier communication networks which typically are structured according to the service they provide. In other words, in the case of Internet the medium is composed of the universal network. It provides the performance-shaping background for practices, even though typically people are not aware of this background effect. It has been indicated "Without the artists' intervention man merely adapts to his technologies and become their servo-mechanisms." ([R23], p. 98). A further typical feature of a genuine medium is that it swallows former media and makes them as its content. This precondition is very clear in IP network that is capable of making use of text, telephone, TV, movies and MP3. Hence it is evident that IP network has a *medium character*. At the same time it provides basis for services that make sense to different contextually defined activities that the IP network may support. In this sense IP also has the *tool character*.

The most important affordance of Internet is it global *connectivity.* Due to connectivity Internet becomes an extension of the human nervous system. If we can assume connectivity being such a generic intrinsic feature we should try to foresee what kind of processes would take place in the implementation of IP network into the daily lives of people, and what kind of abilities people should develop in order to grasp and make use of the affordance of connectivity as fully as possible. We propose the application of the method McLuhan labelled the Laws of media [R23]. The laws of media is a conceptual way of testing and falsifying "observations about the structure and nature of things man makes and does" ([R23], p. 3). The method does not rely on a theory of technology but provides a heuristic but systematic device to understand the nature of media. This device is a tetrad that raises four questions that can be put on any technology or medium. The questions are indicated in the tetrad in Table 2.

By using the tetrad to IP medium a following first approximation of the processes of IP implementation were revealed. These are depicted in the following Table 2. Each point could be concretised by analysing how they are portrayed in peoples' daily practices and in problems that people appear to face in handling their work, leisure or other situations.

Table 2. Sketching the transformations of practices in the age of IP networks using the tetraed method of Laws of Media [R23].

ENHANCEMENT: What does the medium enhance or intensify? Extension of the human nervous system to a global nervous system; being connected to everywhere Significance of tactile relationships, resonance?	REVERSAL: What does the medium produce or become when pressed to an extreme? Paralysis due to loosing focus and presence
RETRIEVAL: What does the medium retrieve that was previously obsolete? Presence in local village Integration of senses	OBSOLENCE: What does the medium render obsolete or displace? Specificity, differences and hierarchy of media? Dominance of the visual space?

IP Connectivity and Dependability of Service

So far we have analysed the generic feature of connectivity and what kind of performance shaping potential this affordance provides for human usage. In this section we shall change the perspective and conceptualise the relationship of connectivity with the dependability of IP medium. In other words we would like to raise the question of how people experience the threat of loosing connectivity.

An intrinsic part of a user's relation to networking and the network-mediated services and functions is the *dependability* of the IP medium, i.e., the questions about its availability, reliability, controllability, vulnerability and security. Can the user rely on this new infrastructure as much as on the traditional technologies and the old ways of acting? High level of dependability should be a natural requirement for a global information infrastructure, but, so far, this has not been the trademark of the Internet. Rather, although the dominant experience is the impressive success of the Internet, its users are accustomed to many kinds of negative surprises.

The traditional telephone network is a highly reliable and stable "machine". It has existed more than hundred years, and it has been rooted in people's consciousness as well as in legislation. There are generally accepted principles, laws and regulations on aspects like confidentiality, and emergency telephony. Now, when the IP medium is approaching the status of the dominant information transmission infrastructure, it is important to recognize the fact that new networks and network-based services emerge and are developed largely in a bottom-up fashion. Cheaper and more effective solutions are rapidly replacing older ones, and this development is faster than the development of regulations that would adequately secure the availability of the new medium in any non-typical circumstances. Part of these questions are already receiving considerable attention, like information security and infrastructure protection, whereas an adequate conceptualization of the dependability of the IP network from an everyday user's point of view and the development of a related measurement and regulation apparatus have hardly entered serious discussion.

The availability of the network layer service in the access network, that is, the successful transmission of IP packets between the user's device and an IP core network, is a prerequisite for all networked applications, and its unexpected failures are often very frustrating for the user. Moreover, the nature and reason of the failure remain usually completely obscure. Although the network can mediate direct criminal threats to the user's safety and property, making the concern about information security more than justified, the DAIMON project wanted to pay attention to less dramatic but still highly significant aspects of dependability. Hours wasted because of lack of everyday dependability of the network medium are paid dearly both in dissatisfaction and economically.

The user's experiences on the dependability of the network focus on *network availability*. Experience of availability is closely related to the experienced Quality of Service. Indeed, if the QoS falls below a certain level, the service cannot be considered available, and the notion of network availability could be extended to the quantified notion "availability for transmission at rate *r*". However, a user of the IP medium should learn to distinguish between unavailability of a particular service and the unavailability of IP connectivity. If only the first is the case, a user may, if he can, try with another application, or with another access network afforded through the multi-access terminal device. Flexibility and innovativeness in the usage of new communication affordances compensate part of the existing lack of dependability and should be taken into account when assessing network dependability as a whole. The significance of the dependability question is differentiated according to age groups, interests, education etc. This may be one reason why it has been to some extent neglected: we live still in a pioneering era of the IP medium, pioneers are not intimidated by temporary difficulties and, on the other hand, the legacy systems are still widely in use and available when IP connectivity fails.

The above discussion concerning how users would consider the dependability of the IP connectivity reveals an issue that we see very significant for users' apprehension of IP affordances in general. It appeared that in order to realistically identify how to react to the fluctuations of the quality of service due to fluctuations in the IP availability, the users should distinguish between the network and the service application. From earlier discussion in the chapter we know however, that user experience is typically focused on the application and users are not able to identify which part of the whole IP medium is relevant to the experience. On the basis of this discussion a following generic hypothesis can be formulated concerning the comprehension of IP affordances:

In order to make use of the major affordance of IP, i.e. connectivity, it is necessary that the user, who basically is interested in making use of Internet-mediated end-services, has to distinguish between the internet (and its access technologies) and the higher level end- services.

Making this distinction between the network affordances and the service functionalities is a sign of network literacy. In contrast to many popular visions, we maintain that the network will not and should not disappear from the users' awareness.

Network literacy is the user's ability to apprehend or "read" the affordances of the IP-network in the environment, and to interpret the quality of network service in a realistic way. This would facilitate appropriate and meaningful usage of available end-services.

A distinction between the Internet medium and the services is assumed to promote a realistic perception of the quality of service (QoS) and the ability to utilise the diversity of the network to improve the QoS. Because awareness of the net/service distinction also promotes awareness of the state and functioning of the network, it would prevent from false blaming of the net or miss-diagnosing of loss of the net.

Empirical studies would be needed to find out what kind of awareness the users should develop of the IP affordances (IP as generic medium) while they focus on the different functionalities of the service that are related to their work, learning or any everyday activities (IP as a tool).

Network Literacy

We introduce the term "Network Literacy" to indicate the user's ability to apprehend the connectivity affordance of the IP network in the environment. Grasping the affordances facilitates appropriate and meaningful use of services. By naming this ability as "literacy" we refer to a particular type of relationship in which the interaction between human and environment is mediated predominantly via signs that distantly denote the actual real world. Technology is in this connection experienced as part of the environment, because people perceive the signs that denote the environment, not (predominantly) the actual world. It is foreseen that, in the future, intelligent technologies are embedded in the environment. IP network is assumed to create continuity to perceptions of the environment, to enable communicating perceptions, and to facilitate collaborative activities in the environment. This would result in an even more concrete merging of technology with the environment. Technologically augmented reality creates new affordances and, hence, new practices for grasping these must be created.

The concept of “network literacy” is not very widely used so far. It appears that it is typically used as an extension of the “computer literacy” which has been defined and used already since 1970’s. It has developed into a particular subject in computer pedagogic aiming at teaching people skills to use computers. Hoffman and Blake give an instructive overview of how the concept and content of computer literacy has developed [R13]. One definition of network literacy that we could find for it in the literature indicates that “Network literacy describes the capacity to use electronic networks to access and create resources, and to use electronic media to communicate with others” [R8]. This definition was adopted by the author from the British Department for Education and Employment. No further determinants are mentioned by the author that would relate to the particular nature of IP medium.

Computer literacy/fluency has usually been defined as particular skills that are acquired in formal education at schools and universities. Recently, changes have appeared in the acquisition of computer literacy that might necessitate reconsidering the whole conception of computer literacy. Hoffman and Blake crystallise these changes into three points [R13]: The first point is that children and students in today’s schools and universities are quite skilful in computer and web usage already when they enter the computer classes. Secondly, learning takes place spontaneously and the content of skills is specified according to the students’ and users’ own demands. The authors raise the question whether informal learning would in general be more effective in acquiring computer literacy in today’s situation. They also state that “students learn about the technology if they can relate it to their lives” ([R13], p. 222).

Rückriem provides results of studies on learning computer literacy at university level that cohere with the above authors' views. He shows that the acquisition of skills was much more effective when students first had constructed a societal context and personally meaningful perspective for the usage [R28]. Rückriem does, however, not contrast formal and informal learning but makes the point that an activity and communication system is needed to adopt a medium into wide societal usage. Literacy in its traditional sense was adopted in the school system. If the schools are not functioning in this mission appropriately, what would be the activity and communication system within which the uses of computers and IP network will be adopted and culturally mediated?

A third point that Hoffman and Blake made was that even though people are well equipped with using the computer and Internet they do not have knowledge of the technological or other scientific foundations of the media they use.

On the basis of above reasoning and in connection to the hypothesis proposed in the previous section we would like to formulate a second hypothesis:

Depending on the level of network literacy different societal perspectives to the implementation of IP-medium will be opened:

A) The consumer perspective: Merging the network with end-services leads to adoption of the role of a service consumer, i.e. forming an audience.

B) The authorship perspective: Identifying the Network as a generic medium leads to active usage and authorship, i.e. forming a public.

We propose that creating network literacy and facilitating the authorship perspective to IP would assume at least the following five preconditions. Some important issues under each point are also listed. These should be considered as questions for further research.

1) Forming a personal sense and perspective for the usage of the network

2) Developing new practices and competencies to grasp or apprehend the IP connectivity

   Capability of apprehending the intrinsic features of IP would promote:

   – coping with complexity and uncertainties: IP-media provides a new catalogue function, finding innumerable perspectives to the environment and the world
   – coping with simultaneity and instantaneousity [R22]: reference to past and future, personal lifespan
   – coping with global presence: embodiment, and locality is needed for creativity and identity, ability to act ([R2], [R4])
   – coping with interactions; responsibilities, habits, norms, values.

3) Creating trust in the medium

   – Medium is external part of the human – there might be lack of trust in oneself when using own creations
   – Understanding the constraints for the use of the medium.

4) Awareness of the different forms of governance of the network

   – Different scenarios have been sketched ([R4], [R7], [R30]): users as consumers or users as authors.

5) Involvement in a communication system

   – Meaningful usage of the medium develops in activity systems [R28], what are these?

As can be inferred from the above our conception of network literacy describes an ability to use a new type of medium for interacting with the environment. Network literacy is not a mere extension of the literacy conception which is based on written alphabetic text. It could even be proposed that instead of speaking of literacy we could use the notion of *teleactiveness* in order to emphasise that “reading” the environment is grasping, i.e. perceiving and acting upon the distant environment in a multimodal way.

5.3.2 Network Perspective

5.3.2.1 Existing Wireless Access Networks

Community wireless broadband efforts have been made by applying different models. On one hand using a public, exclusive model, in which the city builds, operates and sells or gives access using taxpayers money or as an opposite using a private exclusive model, in which a single company builds a wide area network and sells paid access or free ad-supported access to residents, visitors and even the city.

Cities and communities in Finland providing large scale WLAN networks include Mäntsälä (MSOYnet network), Oulu (panOULU network), Turku (SparkNet and OpenSpark), and Lahti (Mastonet). The first initiators to build such networks were energy companies broadening their product assortment. These commercial networks were designed by professional network planners and after demand updated according to current standards and practices, e.g. by adding pre-WiMax links for backhaul, QoS and full mobility. The other approach in Finland has been made by higher educational establishments sometimes in conjunction with the cities or private organisations. The set of users is in most cases restricted to students or citizens through authentication whereas other networks are open for everyone, and will therefore in the future be regulated by the communications market act. This cooperative approach makes the network patchy and thus more practical for nomadic than mobile usage.

In the USA, the primary reason to build municipal wireless networks for the public has been to overcome digital divide by providing affordable or free Internet access. The amount of larger networks, city hot-zones and city- or countywide networks purely for public safety employment is increasing. Pioneers in city-wide WLAN have been forced to plan again when not fulfilling e.g. the capacity requirements while new networks believe in offering ad-supported free access, which will include user positioning.

In the classic WLAN network topology, fibre is used for backhaul, since it will last practically for ever and is fast. However, since wired backhaul can be very expensive for establishing wide area WLAN coverage, different wireless solutions have emerged. WLAN with omni-directional antennas as well as point-to-point and point-to-multipoint microwave links and WiMAX have been used. The most common approach for wide area municipal networks is however a mesh-network topology. Airspan, though, says that their experiences in Europe show that cellular, sector based network deployment are technically better than a mesh networking approach.

In a mesh network (Figure 3), access points can communicate with each other without being routed through a central switch point. Network costs depend on the geography and demography, number of users and usage peaks. The IEEE 802.11 standard does not include a non-infrastructure mesh topology. As long as there is no finalized standard, all vendors have their own solutions, e.g. for self-configuration of nodes, traffic signal hopping and routing and they are not interoperable with each other. Vendors have applied both bridged and routed mesh networking where the backhaul either consists of a shared network or a series of point-to-point links.

There exists also WLAN mesh systems focused on sensor and telemetry applications. One system especially designed for public safety applications, but also supporting 802.11a/b/g besides the 4.9 GHz spectrum does not rely on the fixed access points but on the mobiles themselves that by the help of software can act as nodes in an ad-hoc mesh network.

Most WLAN mesh equipments in outdoor citywide networks have nowadays multiple radios, such as WLAN, WiMax and 4.9 GHz. Having more radios and thus supporting several systems in one single gateway increases the cost of the equipment. Some vendors support QoS and security, especially since VoIP is becoming a demand in municipal WLAN networks. WiMax is still much more expensive than WLAN, but since they work on longer distances they can be more suitable for rural wide-area coverage for backhaul. Pre-IEEE 802.11n solutions, i.e. access points utilising MIMO (Multiple Input, Multiple Output), have also appeared on the municipal WLAN market. Satellite connection combined with GSM and 3G networks has also been used as backhaul for WLAN client access for example on board trains and airplanes.

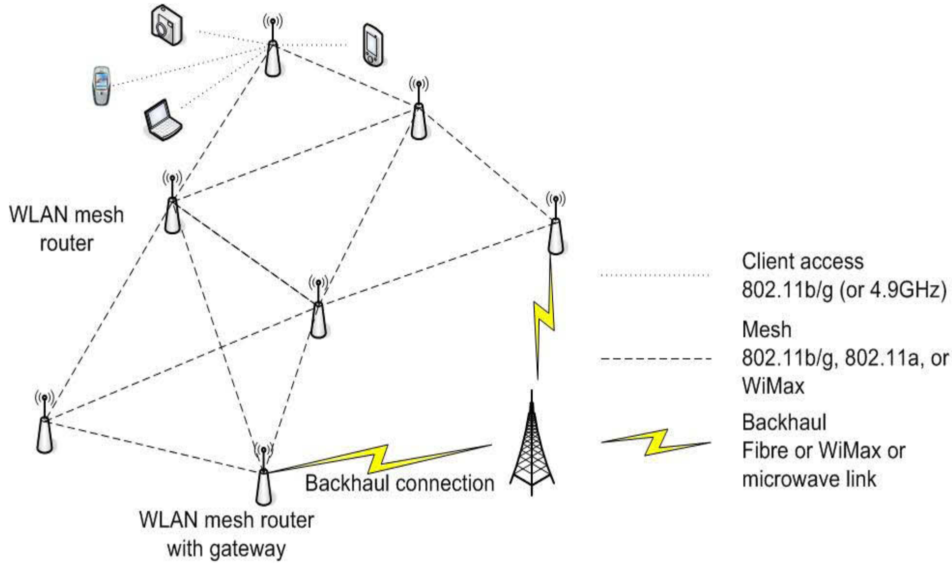


Figure 3. Basic WLAN mesh network structure.

5.3.2.2 DAIMON Architecture

State-of-the-Art

The present day telecommunications networks are a combination of circuit switching and packet switching networks, but as was stated in the introduction (Chapter 5.1) the circuit switching is gradually giving way to packet switching. During the transitional period both techniques live side by side. Examples of this symbiosis are present in 2.5G (GSM/GPR) and 3G (UMTS) networks, both supporting circuit and packet switching (see Figure 4).

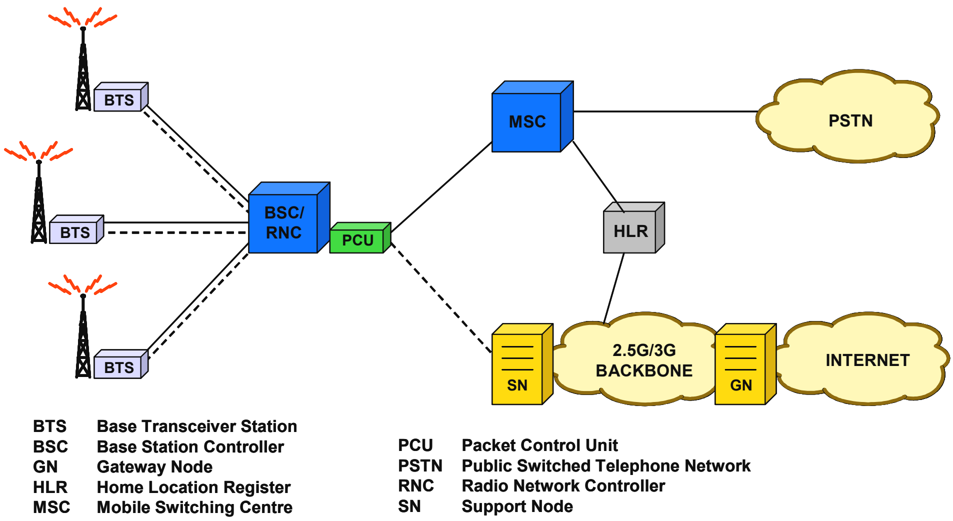


Figure 4. Simplified 2.5G/3G network architecture.

As the evolution proceeds, circuit switched networks shrink to isolated islands, which are connected together via packet based networks by emulating the circuit switched connections. Connection control mechanisms of the circuit and packet switched parts of the network are separated. Some operators have already started to carry PDH circuits (e.g. E1 or T1) over their packet switched backbone network.

The circuits can also be terminated at the edge of the packet network. The carried data is converted to the form supported by the packet network and the converted information is carried through the packet network over a packet switched connection. Connection control mechanisms have to work together, implying additional complexity. Interworking of the conventional telephony service and VoIP service is a good example of this sort of a solution.

In both of the above cases, the packet switching network acts as the core network, but in the latter case the packet core integrates the different access networks (circuit and packet switching ones) to a common IP based service network/infrastructure as shown in Figure 5.

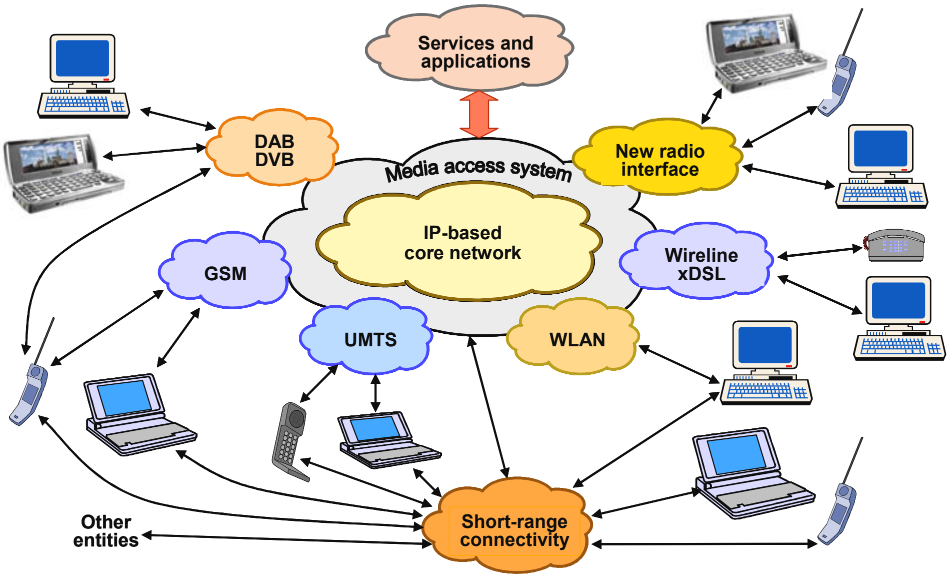


Figure 5. All-IP core network as an integrator.

Starting Point for Architecture Study

ITU-T and 3GPP are developing the future network concept that is based on packet switching. Due to the legacy burden, the future packet network should support also the old established circuit switched networks. The proposed concepts are very complicate and are getting even more complicate as the IP based networks should support the quality and reliability of the circuit switched networks. IP networks do not inherently support circuit switching quality and adding those features means that new protocols need to be defined.

ITU-T's NGN and 3GPP's concept were considered as the starting point for the study of the all-IP network architecture. The target was to foresee the future development up to the point when the circuit switching has no foothold in the basic network architecture.

All-IP Architecture

The all-IP architecture, considered as the DAIMON architecture, assumes that all network connections, starting from a user device and ending up to some server or another user device, are based on packet switching. Taking the mobile network as an example, the major changes in the network architecture are that the circuit switching part and related network elements disappear, different access networks are controlled by common network entities and new types of user, access, connection, service and etc. control units (servers) appear into the network (see Figure 6).

Based on the anticipated architecture, we developed scenarios about the functionality present in the future networks. The functionality must reside in some network elements (servers) and one part of the work was to estimate the type of network elements that are needed and where in the network they should reside. As the network develops into an all-IP network, its architectural structure gets less complicated. However, the need to have a carrier grade all-IP network means that additional complexity has to be added to sew up shortages of IP.

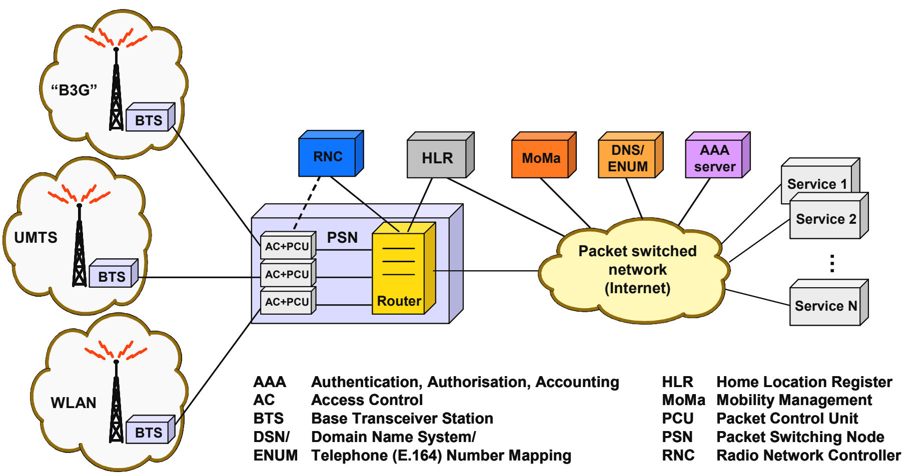


Figure 6. All-IP mobile network architecture.

IP networks are more vulnerable than the conventional telecom networks. In conventional networks, a lot of attention has been paid to reliability of connections and security of communication. For example, key network elements and functions are duplicated and control functions are isolated from the outside world. IP based networks are open in nature meaning that control functions, even the crucial ones, can be accessed from any corner of the network. Additional security mechanisms are needed to protect the network from malicious use.

The triumphal march of the IP technology is also based on lower cost, which is obtained by less secured network elements and connections as well as by less strict technical specifications. Since the IP networks are becoming the corner stones of the future society, it is highly justifiable to study also the dependability issues of the all-IP networks to pin-point the most vulnerable parts of the network.

5.3.2.3 Mobile Backhaul

In wireless networks, the base stations that communicate with the wireless terminals connect to the wire line core network via base station controllers (a.k.a. radio network controllers). Connections, either wire line or microwave, between the base stations and base station controllers form the wireless/mobile backhaul network.

Backhaul Challenge

In existing mobile networks, the backhaul network usually builds on PDH, ATM/PDH, SDH, ATM/SDH and Frame relay/SDH techniques. These techniques offer reliable transfer of user data and control information between the core network and the base stations. These transport concepts are also scalable, i.e. large tree shaped network branches can be built without encountering problems.

As the IP technology advances and the telecommunications market gets more liberal, it has become possible to hire low-cost transport capacity from (local) network operators. As for the mobile networks, the problem is that these operators often sell pure IP or Ethernet transport, which means that the bit pipes coming from the base stations should be converted to IP/Ethernet form or carried over IP or Ethernet in dedicated tunnels.

IP and Ethernet lack many features considered essential for reliable backhaul transport. For example, protection of transport paths against failures and restoration of lost connections (features already well proven in the SDH technique) are missing or imperfectly defined. Scalability is another factor not being in an adequate level in the Ethernet concept. So even though IP and Ethernet are cost-effective solutions, they do not suit for mobile backhaul as such due to the mentioned shortcomings.

Focus in L2VPNs and L3VPN

The starting point for the mobile backhaul study was the work done in IETF, IEEE and ITU-T working groups. Some companies, such as Cisco, Juniper and Nortel, have also contributed to the field. The first phase of the project was to clarify the general principles to carry base station load over IP or Ethernet. After that we proposed co-operation to some operators and product manufactures and finally launched a collateral project, which concentrated on studying commercial products and their technical features. Another project was launched to study dependability features of all-IP networks.

The DAIMON project concentrated on studying technical details of layer 2 (Ethernet) and layer 3 (IP) VPN solutions, related signalling alternatives and their applicability to mobile backhaul. The evaluation of the commercial products approached so close to the DAIMON project that it has been difficult to get in mutual understanding about the result that can be published. At the time of writing this report no publications have been submitted so far in the field of mobile backhaul.

IP/Ethernet based backhaul

The outcome of the backhaul study is summarised in Figure 7. A two-tiered network model was assumed where the BTSs and RNCs attach to a "VPN Backbone" network through "Access Networks" (not to be confused with RAN as a whole.

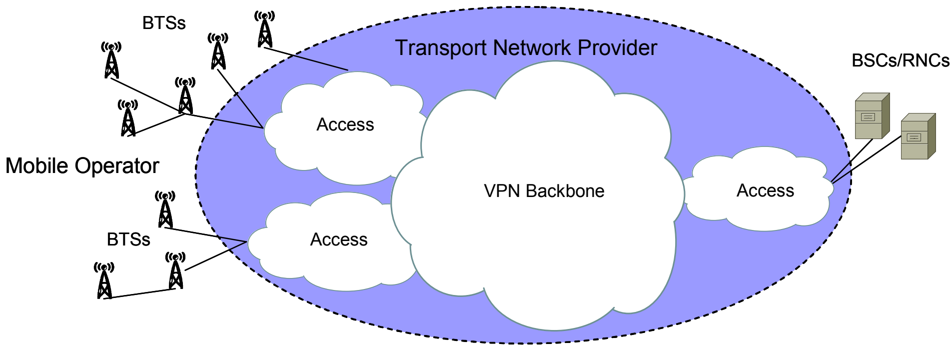


Figure 7. A backhaul network model.

The VPN Backbone provides transport services to multiple separate sets of customer networks each of which can be part of multiple VPNs, although Figure 7 shows only a single RAN. For example, the access networks can interconnect enterprise networks at dispersed geographical locations or they can provide residential/corporate access for connecting to the Internet. A specific feature in the BTS backhaul scenario is that traffic flows only between BTSs and RNCs and never between separate BTSs.

The VPN backbone and the access networks are managed by a VPN service provider that offers transport service to a mobile operator. The mobile operator manages the BTS and RNC sites. Optionally, the mobile operator can lease network access service from an access network operator, which may or may not be the same as the VPN backbone provider.

The type of a VPN service depends on the desired peering relationship between the access network (or end hosts) and the VPN backbone. By definition,

- An L2VPN switches Ethernet frames and uses Spanning Tree Protocol (STP) as the native network formation and resilience mechanism in the access network. The VPN Backbone could be conceived as a huge bridge or hub.
- In an L3VPN, routers in an the access network (or in the BTS/RNC sites) peer with routers at the edge of the VPN backbone. The interconnected customer networks learn routes to each other through routing information carried by routing protocols. The VPN backbone could be conceived as a huge router.

5.3.2.4 Network Dependability

The BETRAFFIC project identified dependability as an under-researched problem area in the vision of an all-IP network, and activity in this area was set forth in DAIMON.

An extended conceptual framework

The notion of *dependability* is usually defined as a collective term with a number of different aspects such as availability, maintainability and reliability. We concluded, however, that it is not sufficient to assess the dependability of the all-IP network using only methods developed for other large infrastructure systems like power plants and the electric network. Among others, Doyle et al. have pointed rightly out in [R5] that the dynamics of the Internet cannot be understood without an adequate theoretical understanding of its multi-layer communication protocol structure. Thanks to them, an IP network is a remarkably self-controlled system. Therefore, we added to the picture of dependability three more aspects with system theoretic flavour:

- the *robustness of design* of the basic algorithms and protocols, meaning stability with respect to arbitrary inputs
- as a counterpart to the former, one should pay attention to *vulnerabilities of design* that often are, in fact, neglected aspects of the generally robust design
- finally, despite the self-controlling character of part of IP technology, we want to include certain amount of *controllability* to the aspects of network dependability.

The BETRAFFIC document presenting this approach to dependability was not published, but its ideas were contained and further developed in the extensive baseline paper of the IPLU project [R1].

Surveys on DNS and BGP

Two elements of the current IP architecture were selected for special analysis because of their doubtful role from a dependability point of view: the Domain Name System (DNS) and the Border Gateway Protocol (BGP).

DNS is not a part of the IP itself but necessary for practical usage of the Internet. It is a huge hierarchical distributed database of domain and host names and their associated IP addresses. Name servers belong to the most popular targets of denial-of-service (DoS) attacks. On the other hand, the vulnerability of DNS has been well recognized, and the threat has so far not fully realised despite of numerous malicious attempts and also, some cases of unintended DNS overloading by bad software implementations. The main countermeasure has been mirroring of root name servers to a large number of hosts around the world. Another aspect, improving the security of DNS, has not advanced as well.

BGP is the protocol responsible for most inter-domain routing in the Internet. Its design is generally considered unsatisfactory for its central role in the architecture. Besides stability problems like route oscillations, it is dangerously sensitive to erroneous information from peer routers. It was also observed that Internet service providers' trend to replace the use of centralized exchange points by bilateral transit agreements poses a risk to the robustness duo to the resulting cyclic topologies.

Dependability Aspects of the DAIMON Architecture

Let us now focus on the DAIMON architecture, leaving the general problems of the Internet architecture aside. Then, the fundamental requirement of the architecture is the provisioning of IP connectivity between the mobile users and the Internet. Further requirements are provisioning of IP connectivity with specified QoS, mobility support with seamless handovers both horizontally and vertically, and the availability of location information for services that need it. Considering the DAIMON architecture from the classical structural reliability point of view, two consequences can be drawn:

First, the provision of several independent radio accesses yields a remarkable increase of the overall dependability of IP connectivity. However, since the default access is typically that with highest transmission rate and quality, the QoS level can seldom be maintained and the default access is lost.

Second, the integrated design where several wireless access networks share the instances of system components like DHCP, DNS and AAA servers poses high dependability requirements for these components. To avoid single points of failure, it is not sufficient to invest in higher quality hardware and software, but it is important also to build more redundancy in the whole system, making the economic gain of the integration somewhat lower than one might think at the first glance.

5.3.2.5 Traffic

One of the main theoretical problems is the lack of generic traffic models in the cases where traffic aggregation level is low. Such models would be valuable tools in planning and dimensioning ubiquitous future networks. One motivation of the studies presented in [P6] and [P15], both started within the DAIMON project, were on such traffic models. Especially, [P6] will pursue in this direction within the framework of the ABI project.

In a preliminary research, made for the study presented in [P7], wavelet analysis was applied to aggregate TCP traffic in order to infer signs of possible congestion in the network. This analysis was based on the assumption that the most common (or dominating) RTT values of the aggregate traffic cause some rough periodicity that is visible in some time scales. While wavelets were found useful for the analysis of an aggregate of TCP flows, ordinary periodogram (Fourier analysis) was found useful for the analysis of an individual TCP flow. The use of these signal processing methods was promising and further research on this topic is planned to continue, possibly within the ABI project.

5.3.2.6 Radio Resource Management in all-IP Networks

UMTS

The resource management theory was presented for UMTS in a project report issued in 2005. Here, part of that report is recalled.

There are two main resources in UMTS that have to be managed: uplink interference and downlink transmission power. Both of them are shared by the users. However, their characteristic is non-linear: the resource allocation increases the faster the higher is the system load. At a certain load level the theoretical interference (in the uplink) or the transmission power (in the downlink) approaches infinity – this load level defines the maximum system capacity. Resource management cannot affect it, but it enables controlling the way the resources are shared.

When the IP-based services are considered, the following resource management problems are most relevant for the UMTS:

- Scheduling and High Speed Downlink Packet Access (HSDPA)
- Priority levels and QoS techniques
- Multimedia Broadcast/Multicast Service (MBMS)
- Unlicensed Mobile Access (UMA).

The scheduling and HSDPA addresses the asymmetry of most of the IP-based services and enables transmission adaptation to the channel conditions (HSDPA). The priority system enables service differentiation and different treatment of the different QoS levels. Finally, the MBMS proposes combining and sharing downlink transmissions of common content (e.g. TV, radio). In this way the resource allocation is minimised. The UMA does not address services directly, but it proposes methods for UMTS and WLAN integrations thus enabling smooth transition from UMTS to the purely all-IP world. One of these topics, namely scheduling of asymmetric services, was further studied in simulations.

The simulations were performed with a dynamic UMTS system level simulator created in CAUTION project (co-financed by the European Commission). All the main UMTS features and procedures (including the power control) have been implemented according to the ETSI specifications.

In order to study the impact of the DL data transmission on the lost capacity of the system, two types of data service (called "heavy" and "light". because they differ from each other in the amount of offered traffic) were implemented over 384, 144 and 64 kb/s UMTS dedicated channels. The idea of outage was based on the UMTS capacity model developed at VTT. In brief, outage occurs when a user whose signal quality is poor cannot be dropped immediately from the network. In order to avoid massive droppings whenever the radio conditions change, such outage users must be allowed to remain in the system, even though their transmissions are ineffective. However, the interference they generate decreases the overall system's capacity. This loss is called "lost capacity" and was analysed in the DAIMON simulations. The amount of lost capacity due to simultaneous data transmission is presented in Figure 8.

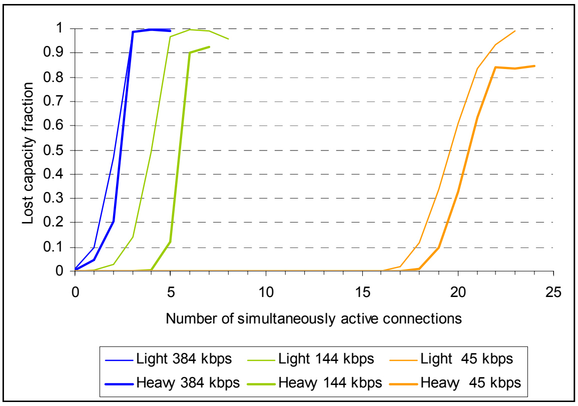
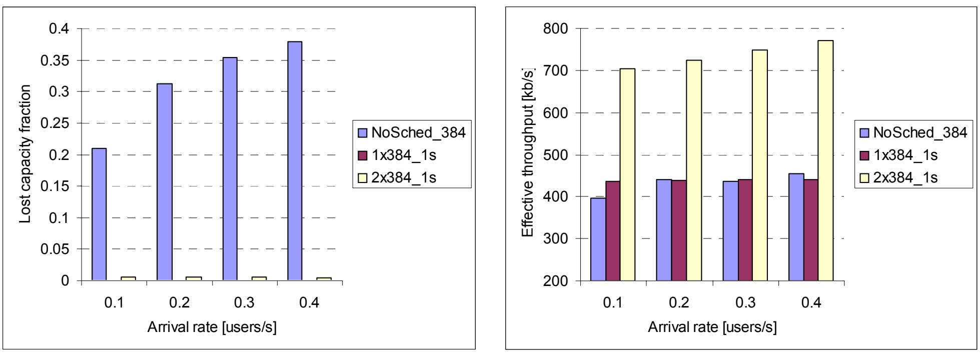


Figure 8. The outage caused by simultaneous data transmission in the DL UMTS radio interface.

The figure clearly indicates the need for scheduling: for every analysed case it is possible to find a limit, up to which the outage is negligible, but beyond which it raises very fast. This limit is the scheduler's input data. To verify the applicability of the resource management technique, the simulator was enhanced and the technique simulated. In this report only selected results are presented: in Figure 9 the lost capacity fraction and the effective throughput of a heavy data service, transmitted over 384 kb/s channel, are presented for 4 arrival rates and 3 cases (no scheduling, scheduling with only 1 active transmission and scheduling with 2 active transmissions). The plots clearly indicate the positive impact of the scheduling. This means utilisation of resource management in UMTS when the IP-based services are considered, is necessary.



(a) *(b)*

Figure 9. Impact of scheduling on the DL lost capacity (a) and effective throughput (b).

WLAN

WLAN has been developed to enable easy and fast network setup as well as distributed access. QoS has not been addressed from the beginning and consequently, much research on this topic can be found. For service differentiation and QoS provision the parameters that control access to the channel are tuned. Moreover, admission control and bandwidth reservation can be used. Adapting to different channel conditions by using link adaptation and power control is also included in radio resource management for WLAN networks.

Service differentiation and QoS is only one objective for radio resource management. Other objectives include fairness between stations. Furthermore, QoS for a specific user might be contradictory to maximising the overall channel utilisation, the bandwidth utilisation and the throughput.

The QoS concept includes controlling the delay, jitter and packet loss in the network in order to support delay-sensitive services (e.g. voice and video) and to allow high-priority traffic frequent access to the channel. In 802.11e the QoS issue has been addressed by introducing and arbitrary interframe spacing (AIFS) and by adjusting the contention window (CW). Using the AIFS parameter and adjusting the contention window improves QoS for high priority traffic but the performance of low priority traffic usually becomes worse. The Hybrid Coordination Function (HCF) uses polling and scheduling which utilises the HCF is also one way of providing QoS.

The literature review showed that in service differentiation and QoS research different radio resource management techniques are combined. Service differentiation can be achieved by:

- adapting the contention window, IFS and the backoff algorithm
- adjusting the transmission opportunities (TXOPs), the service interval or by additional polling.

Under low network load conditions fast access to the channel is enabled by using short contention windows. As the network load increases the length of the contention window should increase. According to the literature, the values of CWmin and CWmax should be changed dynamically and gradually. For service differentiation the arbitrary IFS (AIFS) parameter is more appropriate to use than the contention window. Another way of providing service differentiation is to distribute TXOPs depending on how sensitive the traffic is to delays.

Furthermore, admission control is also included in the QoS concept. Admission control can be based on several different strategies, which aim at maintaining a reasonable QoS for the ongoing connections and providing good enough QoS for the new connection. First, the constraining parameters, e.g. maximum or average bandwidth utilisation, maximum delay, maximum collisions and minimum throughput, have to be identified and second, they have to be measured or modelled. A general survey of admission control schemes in 802.11e networks is presented in [R9]. Both enhanced distributed channel access (EDCA) and HCF controlled channel access (HCCA) are considered. Admission control in EDCA is either measurement-based or model-based. The measurement-based methods admit connections based e.g. on a temporal admission to allow measurement of the throughput and delay. Model-based admission control can be based on models for the saturation throughput. Also in [R9] it is stated that the best solution could be to merge the measurement- and model-based approaches. Admission control for HCCA could be based on the TXOP parameter that corresponds to a certain data rate.

Apart from service differentiation, adapting to different network conditions is also included in QoS provisioning. Service differentiation works well as long as the network is not congested but under high traffic load the following methods can be used:

- admission control and bandwidth reservation
- tuning the IFS and contention window parameters further
- speeding up the backoff countdown when the channel is idle.

Scheduling and admission control are resource management techniques in the MAC layer. However, resources should also be managed in the physical (PHY) layer by using, e.g. link adaptation and power control. Link adaptation is sometimes included in the QoS concept. The 802.11 standard defines multiple transmission rates that are achieved using different modulation techniques, but rate adaptation based on the channel conditions is left open. Higher transmit rates offer larger throughput but they are also more sensitive to channel noise. According to the literature rate adaptation to channel conditions can be done based e.g. on received signal strength (RSS), packet error rate (PER), acknowledgements (ACKs) of transmitted frames and signal to noise ratio (SNR). Rate control can also be divided into statistic-based rate control, which uses e.g. throughput, frame error rate (FER) or retry counting, and SNR-based rate control. Statistic-based approaches respond slowly to channel conditions but SNR-based approaches that are responding faster have to deal with the problem of obtaining reliable SNR measurements.

In 802.11 all packets are transmitted at the maximum power level. A high power level is unnecessary if the path loss is small. In ad hoc networks the power control can enable simultaneous transmissions, which has been shown in the literature. In this technique, stations agree on the minimum transmission power that they require and then stations that do not disturb each other are transmitting simultaneously. According to the literature, power control can also be used to minimise hidden and exposed nodes. Hidden nodes are nodes that might transmit at the same time without being aware of other nodes. Exposed nodes are nodes that could have simultaneous transmissions but they are silenced by the carrier sense mechanism. The transmission power should be set in such a way that the carrier is not sensed at a distance that is larger than the distance from which another node could interfere with the transmission. Furthermore, higher transmission power enables the use of a modulation technique that requires a higher SNR but that is able to transmit more information.

WiMAX

In the development of the WiMAX standard emphasis was put on network performance and QoS providing through collision-free connections and bounded delay. Radio research management in WiMAX is quite a fresh topic, but it can be observed that the focus is on scheduling and admission control, because these algorithms have been left outside the 802.16 standard. Differentiating between the QoS classes and taking their service requirements into account is included in most of the schemes. In the future, handover might be further researched, particularly in Mobile WiMAX. Mobile WiMAX is based on the recent 802.16e standard and it includes support for mobility, e.g. handover and roaming.

Another recent standard is 802.16d (or 802.16-2004), which is known as Fixed WiMAX and it was developed as an alternative to cable and digital subscriber line (DSL) techniques.

The QoS issue has been thoroughly addressed in WiMAX. Four QoS classes are included in 802.16 but Mobile WiMAX (802.16e) includes a fifth class. All QoS classes are listed in Table 3. Except for the ErtPS class the other QoS classes are similar in 802.16 and 802.16e but "Polling" in the name is replaced with "Packet" for 802.16e.

Table 3. QoS classes in 802.16.

QoS Class	Description
Unsolicited Grant Service (UGS)	Real time data streams, e.g. VoIP
Real-Time Polling/Packet Service (rtPS)	Variable-sized packets transmitted periodically, e.g. MPEG.
Extended Real-Time Packet Service (ErtPS)	Voice transfer with activity detection, e.g. VoIP
Non-Real-Time Polling/Packet Service (nrtPS)	Delay tolerant services requiring a minimum rate, e.g. FTP.
Best Effort (BE)	Services that have no requirements.

When scheduling is used for QoS provisioning, three schedulers are usually required for 802.16 TDD in the point-to-multipoint mode. The base station (BS) requires one scheduler for the downlink and another one for allocating bandwidth (TXOPs) in the uplink. The subscriber station (SS) needs one scheduler in order to schedule packets belonging to different service flows. In the literature, different kinds of round robin techniques have been proposed for these schedulers as well as different weighted fair queuing techniques. The literature review showed that UGS connections are usually allocated a fixed amount of bandwidth in accordance with the standard. Within the rtPS class packets with the earliest deadline are sent first. Within the nrtPS class weight fair queuing is often applied and the weight may depend on the ratio between the connection's average data rate and the total average data rate for the whole class. Within the BE class the bandwidth is usually distributed equally among the connections. A cross-layer (MAC and PHY) approach for scheduling is to assign priority not only depending on the service requirements but also depending on the current channel quality and QoS satisfaction. This way, the available bandwidth can be used more effectively. In mesh networks, scheduling aim at enabling simultaneous transmission between nodes that do not interfere with each other.

In order to perform admission control the available bandwidth or capacity has to be estimated and, if possible, reallocated. For this purpose, the QoS class of the new connection has to be taken into account. It can be seen from the literature that UGS connections are usually accepted if the required bandwidth can be provided all the time. rtPS and nrtPS connections are admitted if the required bandwidth is currently unused or if the existing connections can give up some bandwidth. BE connections can easily be admitted since no bandwidth has to be reserved for them. As far as capacity estimations are concerned, admission decisions can be based on the minimum reserved transmission rates for the stations as well as on the number of slots required at the nodes of a mesh network to satisfy the new connection.

The 802.16e standard includes mobility and as this standard gains popularity, handover research might increase. Some handover optimisation algorithms have already been presented in the literature. Continuous scanning and association of possible target BSs waste resources and consequently, these processes should only be done for the candidate BSs that are closest in distance. Furthermore, since the packet flow is interrupted during hard handover, novel management messages have been created to enable packet flow before the handover process has been completed. In mesh networks, it has be proposed that the signal to interference plus noise ratio (SINR) and the estimated data rate can be used in the handover process. Mobility can also be supported in the 802.16d standard by modifying some processes and protocols.

5.3.2.7 Positioning in All-IP Networks

First, we evaluated and enhanced a location method called Database Correlation Method (DCM) in heterogeneous network environment ([P4], [P2]). Second, we improved the existing terminal measurement tool WiNK to support collecting data from WLAN besides GSM networks [P10]. In addition, we designed and implemented a test bed utilising location fingerprinting techniques. Different methods for indoor positioning were studied and one of the outcomes was a preliminary feasibility survey of utilising passive UHF RFID tags [P3].

The designed and implemented WLAN positioning system is composed of three stand-alone components (Figure 10): the client running at the mobile terminal, the server agent handling the location requests and other data services, and a GIS monitor to display terminals' current locations. The design and implementation of the server and mobile terminal client program have partly been based on the thesis work by Y. Li [P11].

To evaluate the DCM algorithm, with an additional site survey tool as part of the developed system, reference fingerprints were collected at 84 locations with 5 meter spacing, having 200 samples for each, on the 2nd floor of Digitalo office building in Otaniemi (Espoo, Finland).

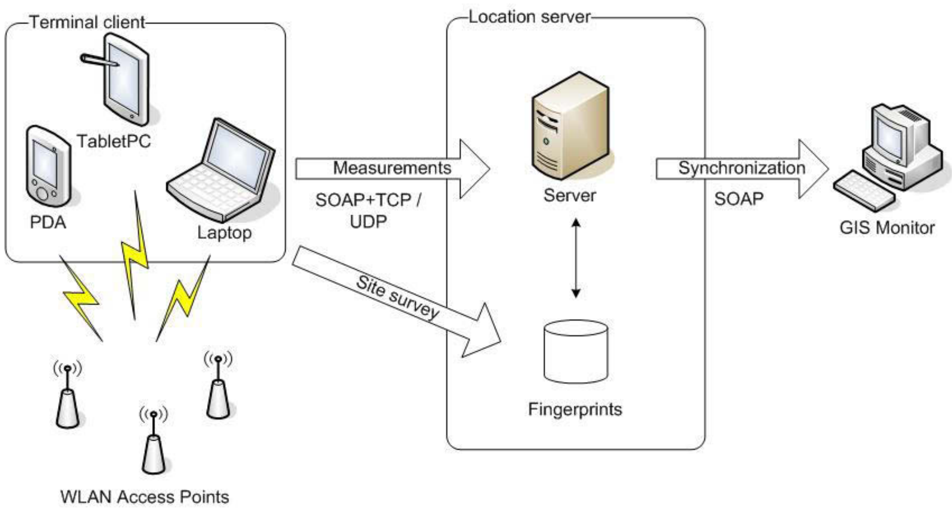


Figure 10. Test bed architecture for WLAN positioning.

A test measurement containing 140 samples was carried out by walking along a planned route (see Figure 11) with a WLAN-enabled mobile terminal. The position estimates calculated with DCM for each of the samples were compared to the real locations in order to verify the location accuracy. In addition, a Kalman filter was used to improve the accuracy.

The cumulative distribution of the achieved accuracy is shown in Figure 12. Two-thirds of the samples are located within 3 meters of the real location. Kalman filtering yields better accuracy: 67th percentile is within 2.5 meters and the maximum error is less than 6 meters.

By increasing the number of coverage-overlapping access points, the positioning accuracy can be improved to a certain degree. However, due to the random variations of RSS utilised in location fingerprinting, positioning techniques like DCM will not be able to provide centimetre-level accuracy.

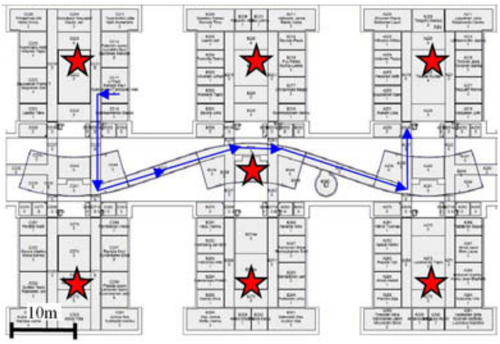


Figure 11. Trial route (blue arrowed line) and access points' locations of WLAN (red stars) on the floorplan of Digitalo office building, floor 2.

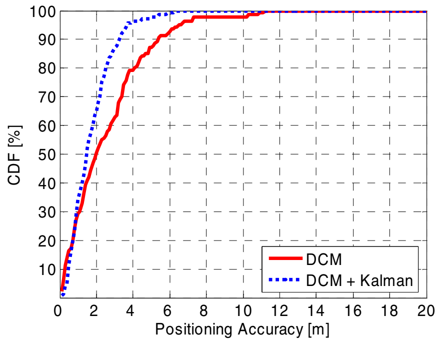


Figure 12. Cumulative distribution of the positioning accuracy of the trial route.

Modifications of the access point placement and in the indoor space (e.g. new walls or large pieces of furniture) will degrade the positioning accuracy if the reference fingerprints are not updated according to the change. Thus, the precision of the positioning estimates needs to be monitored closely in order to be able to detect changes in the propagation environment. On the other hand, site survey measurements and positioning estimates can also be exploited in network planning and resource management. For example, the QoS information together with location estimates can be utilised by the network administrator to adjust the access points to serve the clients more efficiently.

Later on, the implemented test bed will be utilised in the Motive project (European Union) as a part of the platform for ubiquitous terminal assisted positioning, supporting the Secure User Plane Location (SUPL) and OMA Mobile Location Protocol (MLP).

5.4 Conclusions and Impacts

User perspective

As continuation to the work started in DAIMON, we plan to accomplish explorative interviews with different user groups to study the network literacy concept and the exQoS model. Of the latter we are particularly keen to develop relevant categories for appraisal of service (AP) i.e. trustworthiness, sensitivity, engagement, tactfulness, and harmony. On the experience of these interviews we would proceed to the next phase of model building and to the model testing. Thereafter we would enter in a third phase in which development of empirical methods to study peoples' conceptions of IP and IP-based services and practices in using them would take place.

The availability of several access networks increases the dependability of the user's IP connectivity considerably. However, certain network literacy is required from the user to grasp the affordances and to interpret the variations in quality of service.

Preparations are under way to implement the ideas introduced in the DAIMON project into a strategic research project in which the future of all-IP mobile network in the rural areas is studied. The approach that is going to be used in this "Rural Connections" study is the Living Lab conception. According to this approach new technological possibilities are implemented in use within certain activity systems that form new types of intelligent environments. Extensive field studies with intensive user participation are accomplished over a longer period of time. This approach is one of the key instruments that are currently implemented within the European Commission to improve realisation of the information society perspective.

Network perspective

As the transport network evolves to an all-IP network, the network architecture becomes more straightforward and more open. The straightforwardness is caused by the fact that many network elements, necessary in today's combined circuit-packet networks, become obsolete and disappear from the network. The openness is an inherent part of the IP technology enabling all sorts of players to implement their networking solutions and services to the network. The counter part of the all-IP architecture (i.e. the sketched DAIMON architecture) is that it poses high dependability requirements for the system components.

The mobile backhaul network evolves from the conventional PDH/SDH transport towards less expensive solutions, i.e. the Ethernet or IP transport. Both these lack many important features, such as protection of transport paths, restoration of lost connections and efficient QoS supported provisioning of transport capacity, commonly available in the conventional transport networks. The advanced layer 2 and layer 3 VPN solutions seem to be the answer to these shortcomings.

The mobile all-IP architecture study helped to identify relevant development topics and problems in the all-IP networks and to direct our research work towards substantial problems, such as reliability of all-IP networks and problems in the mobile backhaul network. Discussions with the industry and network operators have clarified the picture even more and some follow-up projects have already been launched, one dealing with the dependability evaluation methods for all-IP networks (http://iplu.vtt.fi) and another dealing with the mobile backhaul problems. A co-funded project to study and develop solution for OAM and provisioning of transport capacity in the mobile all-IP networks is under discussion. Additionally, we are involved in preparing a couple of project initiatives to EU's 7th Research Framework Programme.

The evolution of commercial WLAN wide area deployments does not go hand in hand with ongoing standardisation efforts. The importance of mesh networks is growing especially in urban areas. Future all-IP networks were found to require more planning and maintenance. The possibilities for including e.g. QoS are yet rarely used in real implementation.

Besides providing wireless network connections, the existing WLAN infrastructure can offer an inexpensive and quite easily deployable solution for indoor positioning. The received signal information and position estimates can also be utilised in network planning and resource management.

UMTS resource management research was focused on scheduling, QoS provisioning and system integration with other access technologies. In the project an example scenario with a heavy downlink transmission was simulated and analysed to find possible traps and methods to avoid them. The project results clearly indicate the positive effect of scheduling.

In WLANs QoS has not been addressed from the beginning and consequently, much research on this topic can be found. For service differentiation and QoS provision the parameters that control access to the channel are tuned. Moreover, admission control and bandwidth reservation can be used. Adapting to different channel conditions by using link adaptation and power control is also included in the radio resource management of WLAN networks.

The focus in WiMAX is on scheduling and admission control. Most of the proposed solutions consider the QoS classes and their service requirements. In the future, handover might be studied further, particularly in Mobile WiMAX.

The implemented test bed for WLAN indoor positioning will be utilised and expanded in IST-Motive project as a part of the platform for ubiquitous terminal assisted positioning. The analysis of the resource management in all-IP networks and their existing implementations are planned to be utilised in joint projects with industrial partners (concerning network planning and monitoring in heterogeneous systems). The collected information enabled preparing a proposal for the Tekes Ubicom call. The proposed research idea addresses mutual interactions between various wireless technologies that coexist in a modern home or an office (mutual interference, common QoS framework, inter-system routing). The main focus will be on interference between UMTS, WLAN and WiMAX, the technologies that have been analysed in DAIMON.

Publications

[P1] W. Backman, K. Kordybach, “A review of resource management techniques in WLAN and WiMAX systems for QoS provisioning.” European Journal of Operational Research – Special Issue on Performance Evaluation of QoS-aware Heterogeneous Systems. (Submitted)

[P2] P. Kemppi, “Database Correlation for Multi-System Location.” Master’s thesis, Dept. of Computer Science and Engineering, Helsinki University of Technology, August 2005.

[P3] P. Kemppi, “Esiselvitys RFID tunnisteiden käytöstä sisätilapaikannuksessa.” VTT, Research Report, December 2005.

[P4] P. Kemppi, S. Nousiainen, “Database Correlation Method for Multi-System Positioning.” In: Proceedings of IEEE 63rd Vehicular Technology Society Conference, Melbourne Australia, May 2006.

[P5] J. Kilpi, I. Norros, “Call level traffic analysis of a large ISP.” In: Proceedings of the 13th ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management, Monterey, California, USA, September 18–20, 2000.

[P6] J. Kilpi, P. Lassila, L. Muscariello, ”On the dependence of Internet flow traffic.” Extended abstract in the Second EuroNGI Workshop on “New Trends in Modelling, Quantitative Methods and Measurements” (WP IA.8.1), Aveiro, Portugal, November 24–25, 2005.

[P7] J. Kilpi, P. Lassila, “Micro- and macroscopic analysis of RTT variability in GPRS and UMTS networks.” In: F. Boavida et al. (Eds.), NETWORKING 2006, LNCS 3976, 2006. Pp. 1176–1181.

[P8] K. Kordybach, P. Raatikainen, “Data traffic, scheduling and lost capacity in UMTS downlink radio interface.” Proceedings of 13th International Conference on Telecommunications, ICT, May 2006.

[P9] S. Lallukka, P. Raatikainen, “Passive Optical Networks – Transport concepts.” VTT Publications 597, Espoo, 2006, 123 p. ISBN 951-38-6706-4; 951-38-6707-2.
http://virtual.vtt.fi/inf/pdf/publications/2006/P597.pdf.

[P10] A. Lattunen, J. Pajunen, P. Kemppi, Y. Li, “Measurement Tool for Multi-Network Fingerprint based Positioning.” Proceedings of IST Mobile and Wireless Communication Summit, Myconos, Greece, June 2006.

[P11] Y. Li, “A Database-driven XML-based Application on mobile Devices.” Master’s thesis, Dept. of Computer Science and Engineering, Helsinki University of Technology, May 2005.

[P12] L. Norros, “Systems Usability: A concept to support the design of smart objects and environments.” HOT Topics article, Carlton University, Human Oriented Technology Lab, Canada. October 2006.
http://www.carleton.ca/hotlab/hottopics/Articles/October2005-SystemsUsabil.

[P13] L. Norros, “Improving design by understanding usage – outline of an ecological design concept for smart environments.” European Science Foundation Meeting on Nanosciences and the future of information technology, Paris, April 4–7, 2005.

[P14] L. Norros, “On the dependability of Internet understood as a medium – A social science perspective.” Paper at the International multi-disciplinary workshop on “Dependability of all-IP networks.” Espoo, May 18–19, 2006.

[P15] Markovich, Kilpi, “Bivariate statistical analysis of TCP flow sizes and durations.” Extended abstract in the EuroNGI Workshop on “Stochastic Performance Models for Resource Allocation in Communication Systems.” CWI, Amsterdam, November 8–10, 2006.

[P16] P. Raatikainen, “On development of networking technologies and pervasive services.” Wireless Personal Communications, Vol. 33, No. 3–4, 2005, pp. 261–269.

[P17] M. Sipilä et al., “Topics in 4G networks and hardware.” Proceedings of the International Workshop on Optical and Electronic Device Technology for Access Network. Aalborg, Aalborg University, Keio University. Tokyo, Japan, 2005. Pp. 5–10.

[P18] T. Suihko, K. Ahola, P. Raatikainen, K. Seppänen, ”VPNs for BTS Backhaul.” VTT Research Report, 2006. 80 p.

[P19] R. Lemminkäinen, ”DNS report.” VTT Research Report, 2005. 20 p.

[P20] M. Alutoin, “BGP report, VTT Research Report, 2005. 15 p.

[P21] I. Norros, P. Raatikainen, “On the concept of dependability of Internet.” VTT Research Report, 2005. 15 p.

References

[R1] K. Ahola, E. Myötyri, I. Norros, L. Norros, U. Pulkkinen, P. Raatikainen, T. Suihko, “The dependability of an IP network – what is it? A baseline paper of the IPLU project.” VTT, May 16, 2006. Available at http://iplu.vtt.fi/iplu_baseline_2006.pdf.

[R2] G. Bachelard, “La poétique de l’espace.” Presses Universitaires de France, 1957.

[R3] F. Bader, C. Pinart, C. Christophi, E. Tsiakkouri, I. Ganchev, V. Friderikos, C. Bohoris, L. Correia, L. Ferreira, “User-centric analysis of perceived QoS in 4G IP mobile/wireless networks.” 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC’2003), Beijing, China, Sept. 7–10, 2003.

[R4] J. D. Bolter, “Identity.” Unspun. Key concepts for understanding the Word Wide Web. T. Swiss, New York, New York University Press, 2000. Pp. 17–29.

[R5] J. C. Doyle, J. Carlson, S. H. Low, F. Paganini, G. Vinnicombe, W. Willinger, J. Hickey, P. Parrilo, L. Vandenberghe, "Robustness and the Internet: Theoretical Foundations." Draft. March 2002. http://netlab.caltech.edu/internet.

[R6] M. Castells, M. Fernandez-Ardevol, J. L. Qiu, A. Sey, "The mobile communication society: A cross-cultural analysis of available evidence on the social uses of wireless communication technology." The International Workshop on Wireless Communication Policies and Prospects: A Global Perspective, the Annenberg School for Communication, University of Southern California, Los Angeles, 2004.

[R7] S. Cubitt, "Multimedia." Unspun. Key concepts for understanding the Word Wide Web. T. Swiss. New York University Press, 2000. Pp. 162–186.

[R8] L. Dawes, "First connections: teachers and the National Grid for Learning." Computers and Education (33), 1999, pp. 235–252.

[R9] D. Gao, J. Cai, K. N. Ngan, "Admission control in IEEE 802.11e wireless LANs." IEEE Network, Vol. 19, Issue 4, July–August 2005, pp. 6–13.

[R10] J. Gibson, "The Ecological approach to visual perception." Boston, Houghton Mifflin, 1979.

[R11] N. Green, "On the move: technology, mobility, and the mediation of social time and space." The Information Society 18:4, 2002, pp. 281–292.

[R12] W. C. Hardy, "QoS. Measurement and evaluation of telecommunications quality of service." Chichester, NY, Wiley, 2001. 246 p. ISBN: 0-471-49957-9.

[R13] M. Hoffman, J. Blake, "Computer literacy: today and tomorrow." Hamden, CT, Consortium for Computing in Small Colleges, 2003, pp. 221–233.

[R14] ITU-T E.800, “Telephone network and ISDN – quality of service, network management and traffic engineering; terms and definitions related to quality of service and network performance including dependability.”

[R15] ITU-T E.860, “Framework of a service level agreement.”

[R16] M. Ivanovich, M. Fitzpatrick, M. Beresford, A. Saliba, “An empirical 3-D user satisfaction model in wireless data networks.” 15th IEEE International Symposium on Personal, Indoor and Mobile Communications, 2004.

[R17] J. Kilpi, I. Norros, “Call Level Traffic Analysis of a Large ISP.” In: Proceedings of the 13th ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management, September 18–20, 2000, Monterey, California, USA, 2000.

[R18] M. Koivisto, A. Urbaczewski, “The Relationship between quality of service perceived and delivered in mobile Internet communications.” Information Systems and e-Business Management 2:4, October 2004, pp. 309–323.

[R19] A. N. Leont’ev, “Activity, Consciousness, and Personality.” Englewood Cliffs, Prentice Hall, 1978.

[R20] R. Ling, L. Haddon, “Mobile telephony, mobility and the co-ordination of everyday life.” In: Katz, J. (Ed.) Machines that Become Us: The Social Context of Personal Communication Technology. New Brunswick, NJ, Transaction Publishers, 2003. Pp. 245–66.

[R21] M. Lister, J. Dovey et al., “New media: A critical introduction.” London and New York, Routledge, 2003.

[R22] T. W. Luke, “Governance.” Unspun. Key concepts of world wide web. T. Swiss, New York, New York University Press, 2000. Pp. 73–87.

[R23] M. McLuhan, E. McLuhan, “Laws of media. The new science.” Toronto, University of Toronto Press, 1988.

[R24] M. Perry, K. O'Hara, A. Sellen, B. Brown, R. Harper, "Dealing with mobility: understanding access anytime, anywhere." ACM Transactions on Computer-Human Interaction, 8(4), 2001, pp. 323–347.

[R25] V. Pietilä, "Kaiken takana on teknologia: Harold Innis ja Marshal McLuhan tekivät välineistä viestintäteoriaa. In: Mediatutkimuksen vaeltava teoria." T. Mörä, I. Salovaara-Moring, S. Valtonen. Helsinki, Gaudeamus, 2004. Pp. 137–164.

[R26] P. Rabardel, P. Beguin, "Instrument mediated activity: from subject development to anthropometric design." Theoretical Issues in Ergonomics Science 6(5), 2005, pp. 429–461.

[R27] G. Rückriem, "Tool or medium? The Meaning of Information and Telecommunication technology to Human Practice. A quest for Systemic Understanding of Activity Theory." Invited talk at Toiminta 2003 – conference held in Kauniainen, Finland in December 2–3, 2003. http://www.iscar.org/fi/, University of Helsinki, Center for Activity Theory and Developmental Work Research.

[R28] G. Rückriem, "Die Erziehungswissehschaft und der Kongress der Pferdekutscher." Bund demokratischer Wissenshcaftlerinnen und Wissenschaftler – Pädagogik +Medien = Medienpädagogik? 2004.

[R29] S. Tamminen, A. Oulasvirta, K. Toiskallio, A. Kankainen, "Understanding mobile contexts." Proceedings of Mobile HCI 2003, LNCS, Springer, 2003. Pp. 17–31.

[R30] M.-L. Viherä, "Digitaalisen arjen viestintä. Miksi, millä ja miten." Helsinki, Edita, 2000.

[R31] L. S. Vygotsky, "Mind in Society. The Development of Higher Psychological Processes." Cambridge, Mass., Harvard University Press, 1978.

[R32] N. Vicari, S. Köhler, "Measuring Internet User Traffic Behavior Dependent on Access Speed." In: Proceedings of the 13th ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management, September 18–20, 2000, Monterey, California, USA, 2000.

6. Autoconfigurative Radio Networks (AuRa)

Tapio Frantti, Marko Jurvansuu, Aarne Mämmelä, Janne Alasalmi, Ville Haataja, Suvi Juurakko, Pekka Koskela, Adrian Kotelba, Heikki Laitinen, Mikko Majanen, Marja Matinmikko, Petri Määttä, Risto Nordman, Marko Palola, Seppo Rantala, Timo Sukuvaara and Ville Typpö

6.1 Introduction

The AuRa project consisted of several parts performed in several units and groups of VTT. The purpose was to collect and utilize cumulated know-how and information from various units of VTT. Publication [P8] puts together state-of-art reports and achieved research results in the AuRa project during the years 2002–2003. The AuRa project was focused on several important topics of wireless communications: spectral efficiency of cellular systems, quality of service issues (QoS) in vertical handovers between the different networking technologies, positioning technologies in indoor and outdoor wireless networks, routing issues in ad hoc networks and graphical user interface (UI) development for wireless networks simulation environment. The purpose of the publication [8] is to help researchers and designers to direct their activities towards better competitiveness in wireless communication area by utilising the expertise of VTT and by increasing their own view on these selected fields. Earlier research work on the field is summarized in the roadmap publication "Communications Technologies – The VTT Roadmaps".

The significance of publication [P8] is based on the carefully chosen research topics that are in line with current international research of wireless community. The radio spectrum is a limited resource and the basic resources and limitations to the QoS fulfilment are bandwidth, transmission power, delay, and the complexity of the system. Therefore, measures for spectral efficiency are important for the comparison of systems. Moreover, if different systems are compared with spectral efficiency, also offered QoS must be included. Section 2 of publication [P8] covers the topic in detail.

Services of the network can be characterized by a QoS. Some of the services are reliable, *i.e.*, they never loose data and mostly they are implemented by having the receiver to acknowledge the receipt of data or by using redundancy or by combining these two methods. However, this produces undesirable overhead and delays. Hence, quality of service can be enhanced in all the communication layers from the physical layer to the applications. These issues are discussed in section 3 of [P8]. That section also covers signalling schemes and QoS parameter mapping in vertical handovers.

Traditionally, in telecommunication networks QoS measurement devices have been used for several years to measure network's quality of service. Currently user experience measurements have steadily increased their interests and the user experience measurements by use of connected terminals to services are required for the advanced QoS clarifications. Section 3 of [P8] covers the developed novel tool for the user experience measurement in QoS clarification, too.

Positioning techniques both in indoor and outdoor environment have also gained extensive interest among operators, product producers, and service providers to offer a new kind of services and to enhance existing services. Section 4 of [P8] offers positioning technologies overview in wireless indoor and outdoor networks.

Mobility of wireless hosts is one of the hottest research topics in the field. Usually, the coverage area is divided into centrally or base station controlled cells providing the terminals connectivity to the central unit, which can be switched to the other cells if required (handover). Another fundamental approach to the mobility of the hosts is independent, self-organising network composed of several terminals equipped with traffic routing properties. These kind of networks are called ad hoc networks. Section 5 of [P8] concentrates on the wireless network routing issues, *i.e.*, ad hoc routing. Section 5 also introduces the developed graphical user interface for the NS2 network simulator.

Section 6 of [P8] concludes the achieved results and present the view of the near future on wireless networking.

Publications

[P1] Frantti T., "Cascaded Fuzzy Congestion Controller for TCP/IP Traffic." Journal of Advanced Computational Intelligence and Intelligent Informatics, Fuji Technology Press, Vol. 9, No. 2, 2005, pp. 92–99.

[P2] R. Van Eijk, A. Salden, A. Peddemors, J. De Heer, P. Määttä, V. Haataja, "Handling heterogeneity in context aware services." International Journal of Pervasive Computing and Communications. Vol. 1 (2005), No. 1, pp. 25–30.

[P3] T. Frantti, "Autonomous Cell Selection Procedure for Cellular Systems." 2nd Int. Conference on Computational Intelligence, Robotics and Autonomous Systems 15–18 December 2003, Singapore.

[P4] H. Ailisto, V. Haataja, V. Kyllönen, M. Lindholm, "Wearable context aware terminal for maintenance personnel." In: 1st European Symposium on Ambient Intelligence (EUSAI 2003). Veldhoven, NL, 3–4 Nov. 2003, Ambient Intelligence. Lecture Notes in Computer Science Vol. 2875. Aarts, Emile et al. (Eds.). Springer-Verlag. Berlin, 2003. Pp. 149–159.

[P5] R. van Eijk, A. Peddemors, J. de Heer, A. Salden, P. Määttä, V. Haataja, "Handling heterogeneity in location information services." In: Communication Networks and Distributed Systems Modeling and Simulation Conference. CNDS 2004. San Diego, California, 18–24 Jan., 2004. 6 p.

[P6] V. Typpö, T. Sukuvaara, M. Jurvansuu, P. Mähönen, "Extending IP micro-mobility to AODV based ad hoc networks." Proceedings of The 1st International Workshop on Wired/Wireless Internet Communications, WWIC 2002, Las Vegas, US, 24–25 Jun, 2002.

[P7] M. Palola, M. Majanen, "Performance of distributed calculation in an active network." The International Conference on Parallel and Distributed Processing Techniques and Applications, PDPTA'02, Las Vegas, Nevada, US, 24–27 Jun, 2002. CSREA Press. Pp. 1056–1062.

[P8] T. Frantti, M. Jurvansuu, A. Mämmelä, “Efficient wireless networking with advanced services and negotiated QoS.” VTT Tiedotteita – Research Notes 2248. Espoo, 2004. 110 p. ISBN 951-38-6473-1. http://virtual.vtt.fi/inf/pdf/tiedotteet/2004/T2248.pdf.

[P9] T. Sukuvaara, “Hierarchical Radio Network Using IP-connections.” Licentiate Thesis, University of Oulu, Department of Electrical Engineering, 2004.

[P10] S. Ahonen, “Laajakaistaisen radiokanavan viivemittauksia hyödyntävät päätelaitepaikannusalgoritmit ja -simuloinnit.” Master’s Thesis, Helsinki University of Technology, Department of Electrical and Communication Engineering, 2002. 76 p.

[P11] T. Perämäki, “Reaaliaikaisen videon siirto UMTS-verkossa hyödyntäen palvelun laatuluokittelua.” Master’s Thesis, Helsinki University of Technology, Department of Electrical and Communication Engineering, 2003. 73 p.

[P12] M. Majanen “Hajautettua tieteellistä laskentaa aktiiviverkossa.” Master’s Thesis, University of Oulu, Department of Physical Sciences, 2003. 76 p.

[P13] T. Frantti, “Estimation of Movement” (I Decrease of Impulse Response Measurement Frequency with the Aid of Acceleration Data, II Speed Estimation via Acceleration Data, III Estimation of Doppler Shift on Portable Terminals Using Acceleration Data). Europe, USA, Japan.

7. Minimum Power Cellular Network (MiPo)

Tapio Frantti, Marja Matinmikko and Aarne Mämmelä

7.1 Introduction

4th generation (4G) cellular systems are recognized as systems that can achieve high-data-rate (100–150 Mbits/s) and spectral efficient (1–10 bits/s/Hz) transmission beyond those currently provided by 2nd and 3rd generation (2G/3G) cellular systems. Existing 3G systems are capable of 384 kbits/s – 2 Mbits/s data rates – 2G systems with enhancements to 56 kbits/s – 384 kbits/s data rates and spectral efficiency of 0.1 bit/s/Hz. The 4G systems will also be capable of delivering multiple services for a variety of applications comparable to those offered by wired networks to users at different locations under very diverse channel conditions. The applications include, e.g., interactive multimedia, network games, and virtual reality applications like telepresence. The 4G systems are also required to support services ubiquitously in different types of environments including indoor, outdoor, and global broadband access. 4G systems should also be compatible with the existing wireless, wired and emerging counterparts. Radio communication sector of International Telecommunication Union (ITU-R) has collected the capabilities of 4G systems in Recommendation M.1645: “Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000”. However, the current bandwidth allocations for 3G systems may not be adequate for the deployment of 4G networks to meet the demand in the years 2015–2020. Spectrum allocation for systems beyond IMT-2000 is under the authority of ITU in World Radiocommunication Conference in 2007 (WRC 2007), which will possibly identify new spectrum allocations for 4G systems. Preparations for the conference are made in Working Party 8F of ITU Radiocommunication sector (ITU-R WP8F), which is responsible for the overall system aspects of IMT-2000 and systems beyond it.

Another main issue involved in the 4G systems development is the choice of multiple access technology to utilize and share available scarce spectrum among a large number of users. The choice will significantly enhance or lower the service quality delivered to users. Existing multiple access technologies used in 2G/3G systems are suitable for voice communications but not for burst data traffic, which would be the dominant portion of the traffic load in 4G systems due to emerging all-IP scheme and high-data-rate requirements. The other important issues in the improvement of the capacity are frequency reuse distance and the effectiveness of the power control. Pico-and microcell structure, where the cell or the sector radius varies from less than 100 m up to 1000 m, with enhanced power control enables effective geographical frequency reuse without excessive in tercell (users in nearby cells cause disturbance) and intra cell (users in the cell interfere with other users in the cell) interference still offering high enough received signal level for reliable signal detection.

The aim of this research was to research capacity improvement of cellular networks with all-IP network architecture that offers ubiquitous high-data-rate transmissions and are capable of delivering multiple services for a variety of applications comparable to those offered by wired net works. The research foci were in the spectrum requirements, efficient power control, and network optimization for microcellular with time division duplex link separation. Challenges arose from the mobility (usually from 5 to 50 km/h) and very high data rate (up to 150 Mbits/s) requirements in indoor and outdoor environment.

The organization of the rest of the chapter is following. In section 7.2 the applied research methods are described. Section 7.3 depicts the developed power control algorithm for the cellular systems. In section 7.4 the power control research is enlarged also for WLAN (Wireless Local Area Networks) type networks due to strong internetworking requirement of the 4G systems. In section 7.5 we consider fundamental stochastic features of the ad hoc network simulations in order to get reliable results for the 4G systems development. In section 7.6 the spectrum requirements of 4G systems and the standardization work ongoing on the field are described as well as VTT's role in it. Finally conclusions are drawn in section 7.7.

7.2 Research Methods

In the research, *literature studies, colloqium lectures* and *numerical simulation* methods were applied. In the literature part, state-of-art reports were created in the earlier phases of the project, and were updated on a regular basis. The researchers also participated actively in international conferences in order to get contacts and to observe the current view of the research community in the field. In addition, researchers participated in the development of standards in ITU. Colloqium lectures given by the leading scientists were used in order to get larger and deeper discussions from the current research topics in the VTT's premise. Radio link and network simulation models were developed in order to obtain a more realistic insight into the system parameters and to get quantitative results.

7.3 Power Control of Cellular Systems[1]

Transmission power control has a significant effect on the capacity improvements in cellular networks. Power control is needed to equalise and minimise the transmitted power of each user equipment (UE) in order to decrease near-far effect and inter and intra-cellular interference. Therefore, power control on the uplink tries to adjust the transmitted power of each UE in the power control area (cell or sector) such that the received power from each of them at the base station is on the same nominal level. In addition to effective stabilization of the received power level, the purpose of the power control algorithm development is to decrease the rise time of the power level from deep fades to a nominal level and to decrease the power level overshoot, downshoot and setting time.

An equalised power level can be achieved with the correct combination of *open-loop* and *closed-loop* power control algorithms. The open-loop power control is used to compensate for the path loss, shadowing and slow fading between the transmitter and receiver ([P1]).

As the UE moves there is a change in the received frequency on each received signal path due to the *Doppler frequency shift*. The Doppler effect causes rapid

---

[1] The section is based on the article Multiphase transfer of control signal for adaptive power control in CDMA systems, Control Engineering Practice, Vol. 14/5, Elsevier Science, 2005, pp. 489–501.

changes (fading) in the received signal level. The faster the UE travels, the more rapid is the fading. However, even a stationary UE may be in a deep fade due to movements between it and the BS. Hence, the received power is the sum of the path loss, shadow fading and fading. The effect of fast fading is decreased by a closed-loop power control scheme.

A. Spectral Efficiency

The capacity of wireless communication systems is dependent on how effectively the system relates traffic capacity to the frequency unit and surface (or volume) element with the used power level, i.e., how *spectrally efficient* the system is. For mathematical definitions of spectral efficiency, see for example [P1]). Therefore, techniques like, FEC (Forward Error Correction) coding (convolutional coding, turbo coding) and error detection coding (block coding, Cyclic Redundancy Check, CRC, coding) as well as ARQ (Automatic Repeat Request), which are required for acceptable error level transmission in theory decrease the payload transmission rate decreasing also spectral efficiency. However, they are necessary for almost error free data transmission to increase payload traffic. Hence, we can argue that the listed techniques also increase the *spectral efficiency* up to a certain limit. The spectral efficiency is also dependent on the required number of radio channels in cells as well as cluster size, reuse distance, sectorization and transmission power level. The net bit rate is also decreased via signalling traffic on control channels and partly on dedicated traffic channels required for different purposes such as paging calls, paging indications, random access for channel reservation, access grants, intra-cellular and inter-cellular handovers, roaming, location updates, transmission of measurement reports and power control symbol transmissions just to mention a few.

B. Clusters, Sectors and Cell Radius

Networking areas of cellular networks are divided to the base station controlled radio cells. Each base station is allowed to use only a limited amount of frequency channels, which can be reused after a sufficient distance (reuse distance) in order to avoid interference from neighbouring cells. Radio cells are combined into clusters and each frequency is used only once inside the cluster but cells in neighbouring clusters can reuse all the frequencies.

In principle, a lower number of cells in a cluster allows a greater number of frequency channels (= higher capacity for the network) per cell. However, a low number of cells in a cluster has also a drawback because it increases co-channel interference due to a short reuse distance. In an FDMA/TDMA network the initial reuse distance can be set to two and the size of a cluster to seven. In theory the reuse distance and the size of clusters of CDMA networks can even be set to one due to the code division of channels with almost uncorrelated Gold codes and m-sequences as well as uncorrelated orthogonal codes. In any case, how ever, even perfect codes autocorrelate and also crosscorrelate to some degree in a real environment, decreasing the signal-to-interference ratio (SIR). Therefore, the size of clusters and the capacity of cellular networks also in CDMA systems are limited according to the required lower bound of SIR. *Sectorization* can be used to increase CDMA system capacity like in TDMA and FDMA systems, too. In TDMA and FDMA systems capacity is increased via frequency spectrum reuse whereas in CDMA systems sectorization decreases interference in the observed cell from the surrounding cells. We should notice, that *e.g.,* the sectorization reduces co-channel interference due to the reduced number of interfering cells, because of directional antennas on base stations instead of omnidirectional isotropic antennas. Hence, the SIR is also increased, but the drawback of sectorisation is the reduced number of channels/sector due to the reallocation of common control and broadcast channels to each sector. In other words sectorisation increases SIR but decreases spectral efficiency and traffic capacity via reduced trunking gain.

A third crucial parameter in radio network planning is the *radius* of cells. With a smaller cell radius a higher UE density is allowed, but practical radio system features dictate the minimum size of a cell due to implementation cost/usage substances. Therefore, in the beginning of a radio network implementation project, the cell sizes are quite large and the size is reduced according to the increased traffic density via cell splitting with new base stations. This approach guarantees an optimal network structure and cell sizes required for effective signalling, handover, quality of service (QoS), roaming and implementation cost versus utilisation.

C. Power Control Schemes

Dedicated traffic channels, i.e., dedicated data channels and dedicated control channels are used to carry user traffic and control messages between the network and an UE in cellular networks. Control messages are also embedded into the dedicated data channels. For instance, in the ETSI/3GPP 3rd generation system, UTRA, the dedicated traffic channels consist of 10 ms frames. Each frame includes 15 slots and each of them has 1 or 2 power control bits. The power control bits from an UE to its BS (or from a BS to an UE) request either a predefined step (1 dB or 2 dB defined by the network) increase or decrease in the BS's (UE's) transmit power.

Here we considered a multiphase transfer of control signals for an adaptive step-size power control algorithm in which the UE adjust its power level in dedicated traffic channels according to the received power control bits (received in multiple, 2 or more, sequential phases/slots) from a base station.

In the presented solution the average power level per space volume can be decreased via power level variance minimization. Therefore the spectral efficiency is also increased. The achieved results with the developed algorithm are compared to a predefined step power control algorithm which is a *de facto* standard of cellular network power control. While this work is mainly motivated by the power control of CDMA based cellular systems and the simulations and results were performed for a UMTS/UTRA 3rd generation system, the approach and the techniques are not limited to these system, but are easily applicable for FDMA (Frequency Division Multiple Access)/TDMA (Time Division Multiple Access) based cellular systems, like GSM. They are also applicable to different kinds of WLAN (Wireless Local Area Networks) systems, which at the moment operate without actual power control. Interested readers can and more details about the uplink and downlink power control of the existing cellular systems in [P1].

D. Adaptive Power Control

The purpose of the adaptive power control algorithm development was to find a general method to effectively stabilize the received power level and to decrease the overshoot, rise time and setting time. The other aim of the adaptive power control algorithm development was to find a method to stabilize the received power level and to decrease overshoot, rise time and setting time without an increase in the control bit rate. Solution of the stated problems is extremely significant for the cellular systems, whose capacities are directly dependent on the interference level.

The developed control algorithm sends a power control command during two or even more control command periods (during the two slots in the case example of an UTRA system presented at the end of this chapter). Hence, the number of commands is enlarged to 16 (four bits) or more (6 bits or more) without an increase in the standardized control bit rate by transmitting control commands in two (or more) sequential phases/time slots. First two bits of command are transmitted in the current time slot and the receiver stores them and waits the next time slot(s), which includes the next/last two bits of control signal. The receiver interprets the delayed control command after receiving all 4 (6 or more) bits. Therefore, the control frequency is decreased to half (third or less) of the original but respectively the dynamic scale is increased to 4 times (or more) the original value (in practise 7 times: 7 values up and 7 down, 1 value for zero and 1 unused) for sign and magnitude binary coding. The power control algorithm can adapt to the changing circumstances much faster than a predefined step power control algorithm with the same control bit rate. [P2]

The adaptive power control system and the fuzzy power controller are presented in detail in [P1]. In a fuzzy controller the input and output variables are represented in a linguistic form after the (normalization and) fuzzification of physical values into linguistic form. The input variables, evaluated from every other received slot, are the power level error and the change of error, and the final output value (encoded to two sequential slots) is the power step increment (two-input, delayed single output control strategy). The power step increment is transmitted during the two sequential time slots: the first two bits in the first time slot and the last two bits in the second time slot with sign and magnitude binary coding. Hence, a receiver waits during one time slot (multiphase delay) after receiving the first bits. The receiver then combines the bits and interprets the contents of them. In the fuzzy reasoning phase, fuzzy rules are presented in a matrix form and Mamdani's inference is used. The rule base is composed by analyzing the dynamics of mobile radio channels' characteristics, transient responses and steady-state properties of the system. The size of the rule base is

25 rules. The control strategy produces the linguistic control output, which is transformed back into the physical domain to find the crisp control output value for the power step. In the defuzzification phase the center of area method (CoA) was used.

E. Results of Power Control Simulation

A simulation model was developed and used to test rigorously the proposed adaptive power control against the existing non-fuzzy control method. In [P3] the principles of the developed simulation model are presented. The MAI (Multiple Access Interference) was here modelled by increasing each user's spread spectrum signal to the system using both short channelisation codes and long scrambling codes.

In the simulations the waveforms of predefined step (±1dB step sizes) and adaptive multiphase controlled signals were compared, respectively, when the velocity was 3 m/s, 10 m/s 20 m/s and 25 m/s and the bit error rate (BER) was 5%, which is a typical value for unprotected transmission in cellular networks. The mobile unit was initially placed in a position which causes a 23 dB path loss. The rise times, overshoot and setting times of the compared algorithms are represented in Table 1. In this comparison one can notice that the rise times, overshoots, setting times and the variances of the fuzzy multiphase adaptive controller are considerable smaller than for the predefined step control system. The rise time for the predefined step controller varies between $13.8–17.3\times10^{-3}$ whereas in the adaptive multiphase controller it varies between $0.6–0.7\times10^{-3}$. Setting times varied in the predefined step control system between $19.0–38.0\times10^{-3}$. In the adaptive multiphase system its values are between $3.2–5.8\times10^{-3}$. Overshoot was around 1.9–5.2 dB for the predefined step controller. In adaptive multiphase controller overshoot stayed at the 0 dB level. The adaptive multiphase controller stabilised the variance level around 1.47–2.29 (6.56–7.31 for the predefined step power controller, see Table 1), which can be explained with the more reactive control algorithm, which can minimize power error. Partly this is due to the predictive nature of the rule base of the fuzzy controller. Waveforms of the comparisons are presented in more detail in [P1].

Table 1. Rise times, setting times, overshoot and variances of the compared algorithms, respectively.

Velocity m/s	1 dB step control				Adaptive multiphase control			
3	0.017	0.028	4.2	7.31	0.0006	0.0058	0.0	2.29
10	0.015	0.038	4.0	6.84	0.0006	0.0082	0.0	1.48
20	0.014	0.023	4.1	6.74	0.0007	0.0042	0.0	1.86
25	0.014	0.019	1.9	6.56	0.0006	0.0032	0.0	1.63

7.4 Power Control of WLAN Type Networks

The power control research was enlarged also for WLAN (Wireless Local Area Networks) type networks due to strong internetworking requirement of the 4G systems. In this work, experimental channel reservation and power control algorithms for IEEE 802.11b WLANs were developed. IEEE 802.11b WLAN network was assumed to present typical WLAN network topology for the coming years. The developed WLAN power control algorithm adjusts transmission power of network stations, which operate in an infrastructure DCF (Distributed Coordination Function) mode. The purpose of the power control algorithms is to maintain received signal power of power-controlled frames on a nominal power level under the influence of channel fluctuations. The power control algorithms are based on absolute power level coding of power control commands. In IEEE 802.11b WLANs control information is conveyed through dedicated control frames or by piggy backing the information into existing frames. The separate control frames for power control are depreciated as they would induce overhead and complicate the transmission of actual data frames. Therefore, the power control commands were piggybacked into CTS (Clear to Send) frames sent by an AP (Access Point) to MSs (Mobile Station) due to fact that interaction between a MS and an AP begins with CTS frame transmission.

The advantage of the developed power control algorithms is considerable energy saving. The disadvantage is slightly decreased network throughput. In addition, a throughput vs. fairness of channel occupancy trade off was observed. Hence, two different kind of channel reservation modifications in connection with power control were also developed for a better overall performance with the power control. These modifications do not induce any overhead.

The performance simulations of the algorithms were con ducted by eight contending stations with free mobility within the simulation area around the AP (radius of 25 m from the AP). The interested reader can find more details about the topic and achieved results from [P4].

7.5 Performance Optimized Ad Hoc Network[2]

A primary aim of ad hoc network performance optimization was to find out fundamental stochastic features of the ad hoc network simulations in order to get reliable results for the 4G systems development. Especially, a concept of convergence time of simulation models was considered. It refers to the minimum simulation time required for the reliable results. In [P5] the concept is explained in detail for interested readers. Its practical numerical model-based evaluation was defined by example simulations. Ad hoc network models considered here were event based discrete time models, which utilize Markov chain theory for the definition of state transitions and Monte Carlo random sampling method for the definition of events probabilities and their duration times.

In stochastic network simulations, numerous reported studies do not contain any information about the convergence time of the model. In some of these publications results are achieved by very short simulation times spanning from a few seconds only to tens of seconds. However, the convergence time even for a moderate size of ad hoc networks, such as 10 nodes, may be thousands of seconds and it increases as a function of nodes as was noticed in this research. Moreover, the same suprisingly short simulation times are used for different sizes of networks despite the fact that larger networks require longer simulation times ([P5]).

---

[2] This section is based on conference paper "Fundamental Features of Ad Hoc Networks' Simulations." The 8th WSEAS International Conference on Mathematical Methods and Computational Techniques in Electrical Engineering (MMACTEE '06), Bucharest, Romania, October 16–18, 2006.

Here an ad hoc network model was used for the evaluation of various network parameters' effects on the throughput. The ad hoc network model was based on IEEE802.11b radio technology and AODV routing protocol. It was noticed that ad hoc network with 10 nodes and pedestrian mobility requires as much as 2000 second simulation time before the throughput converges. A simulation was considered to be converged when capacity fluctuated within 1% limits. Figure 1 presents the throughput for data traffic of the ad hoc network with 10 nodes, AODV based routing and pedestrian mobility as a function of simulation time when simulation time spans from 50 seconds to 5500 seconds [P6]. Here the throughput decreases towards 2000 s but in [P5] communication simulation cases have been shown where the throughput oscillates on the both sides of the convergence value before the convergence time is reached ([P7]).

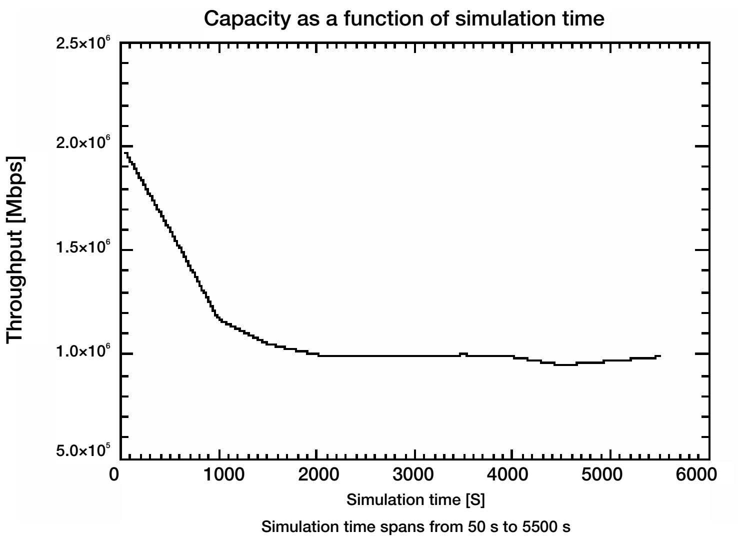


Figure 1. Throughput as a function of simulation time.

In the network model considered here, all the 10 nodes were active. This was modelled so that two nodes (a source-destination pair) were transmitting data (file transfer) and the other nodes were generating background traffic. The background traffic was generated by using random TCP generator and random waypoint algorithm to create node movement patterns in a three dimensional area. All the nodes consumed bandwidth also due to required control traffic for the network topology information update required for the ad hoc networking ([P7]).

The convergence time must be redefined for all the changes in the model, which clearly may have an effect on the capacity (throughput) of the network. For the simulation of *e.g.,* extra transceivers effects the convergence time should be redefined due to an increased number of degrees of freedom. The convergence time has to be redefined also if *e.g,* the routing algorithm is changed or CCK (Complementary Code Keying) modulation/coding is changed to higher performance PBCC (Packet Binary Convolutional Coding) modulation/coding in the physical layer of ad hoc transceivers based on IEEE 802.11b standard or if the connection oriented transport layer protocol is changed to a connectionless protocol. As a rule of thumb, if unsure the redefinition of the convergence time is always preferred and absolutely more acceptable than getting erroneous results with too short simulation times ([P7]).

Figure 2 presents the throughput for data traffic of the same ad hoc network as a function of simulation time when it spans from 10 seconds to 50 seconds ([P6]). It can be noticed that e.g. for capacity simulations if simulation time of 20 seconds, which is quite a typical value in many reported studies, is used an absolutely too optimistic throughput value of more than 3.4 Mbps is achieved. The real capacity is only about 1000 kbps ([P7]).

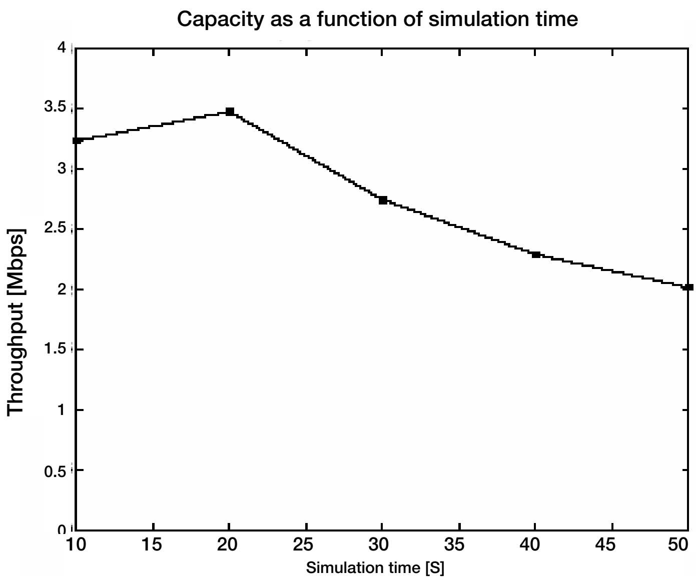


Figure 2. Throughput as a function of simulation time.

From Figure 1 it can be noticed than even if the simulation time is much longer than 20s but still too short, *e.g.,* 100 s, the average throughput of 1.95 Mbit/s for the ad hoc network is achieved. However, the real capacity value is around 1000 kbps, which is about 51% from the erroneous result achieved with too short simulation time ([P7]).

Figure 3 presents an average capacity of an ad hoc network as a function of number of nodes ranging from 3 to 50. For the three nodes the average throughput was around 3.8 Mbps and for the 10 nodes it was only 1000 kbps. The dramatic decrease in capacity was mainly due to an increased number of packet collisions when the number of users grew larger. In larger networks, for instance with 50 nodes, the link failures become even more common and the network is rather congested. The average throughput for 50 nodes is only 116 kbps. This is typical behavior for a wireless systems with a random MAC (Media Access Control) based channel division/reservation algorithm. 116 kbps is very low for a local area ad hoc network with very limited coverage area and only pedestrian level mobility. It is very probable that with the higher level of mobility IEEE 802.11b WLAN standard based ad hoc networks will be fully congested due to increased packet error rate. A throughput value of 116 kbps may be enough for voice communication, poor level video transmission and other low data rate applications such as a short message delivery. It is definitely not enough for the fluent web browsing and email delivery, which are the most common applications of the WLAN IEEE 802.11b front end networks. Therefore, ad hoc networks with higher data rate applications should be based on *e.g.,* IEEE 802.11a technology with admission control, which offers 54 Mbps system level capacity ([P7]).

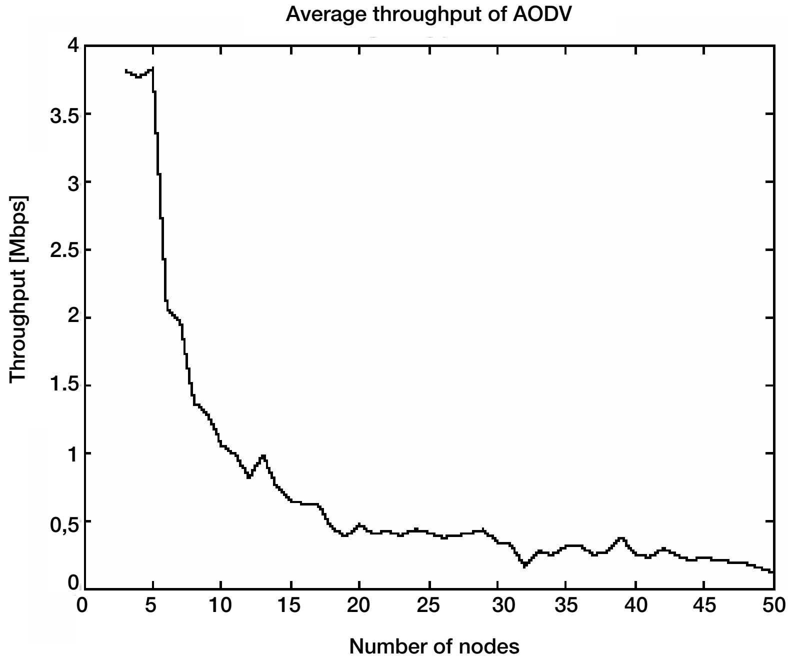


Figure 3. Throughput as a function of nodes.

The *secondary aim* of the research was to optimize IEEE 802.11b AODV (Ad hoc On demand Distance Vector) based network throughput by parameters optimization. In simulations, it was noticed that RTS/CTS (Ready to Send/Clear to Send) threshold had significant influence to the overall throughput in networks, see Table II. The interested reader can find more details about the ad hoc net works and parameter optimization from [P6].

Table 2. Throughput with optimized parameters.

Parameter	**Optimized value**	**Throughput**
Hello interval	Insignificant	1000 kbps
RTS/CTS threshold	3000 bits	2000 kbps
Active route timeout	Insignificant	

7.6 Spectrum Requirements of 4G Systems

A. Role of ITU and WRC

International Telecommunication Union (ITU) is an international organization within United Nations (UN) where governments and the private sector coordinate global telecommunication networks and services. ITU is the internationally recognized entity, which has the sole responsibility to recommend standards for International Mobile Telecommunications (IMT) systems. International Mobile Telecommunications-2000 (IMT-2000), also known as the Third Generation Mobile Systems (3G), is defined by a set of interdependent ITU Recommendations. IMT-2000 provides a framework for world-wide wireless access by a highly capable wireless system of terrestrial and/or satellite based networks. Currently, a lot of effort is globally spent on the future development of IMT-2000 and systems beyond IMT-2000. Systems beyond IMT-2000 are known as IMT-Advanced or 4G systems.

ITU Radiocommunication Sector (ITU-R) plays a vital role in the management of finite natural resources: radio frequency spectrum and satellite orbits. Study Group 8 of ITU-R (ITU-R SG 8) is responsible for mobile, radiodetermination, amateur and satellite related services. Working Party 8F of ITU-R SG 8 (ITU-R WP 8F) is responsible for the terrestrial component of IMT-2000 and systems beyond IMT 2000.

ITU-R WP 8F is the responsible group within ITU for the overall radio system aspects of IMT-2000 and beyond and in particular the terrestrial component of IMT-2000 and beyond. ITU-R WP 8F meets three times a year and develops ITU-R Recommendations and Reports. The framework and overall objectives of the future development of IMT-2000 and IMT-Advanced are described in Recommendation ITU-R M.1645 [R1], which was approved by ITU-R in June 2003.

ITU-R arranges World Radiocommunication Conference (WRC) every two to three years to review and revise the Radio Regulations, which is the international treaty governing the use of radio-frequency spectrum and satellite orbits. WRC is therefore the only authority which can globally decide on the frequency identifications.

In 1992 World Administrative Radiocommunication Conference (WARC 1992) identified spectrum for IMT 2000, which was the starting point for standardization of 3G systems. WRC 2000 identified additional spectrum for IMT-2000. The need for additional spectrum was justified with a methodology ([R2]), which was used to calculate the total spectrum demand of IMT-2000.

Next WRC 2007 will have the following Agenda Item 1.4: to consider frequency-related matters for the future development of IMT 2000 and systems beyond IMT 2000 taking into account the results of ITU R studies in accordance with Resolution 228 (Rev.WRC 03).

Therefore, the next WRC will consider the possible identification of new spectrum to IMT-Advanced (4G) systems. Globally, ITU-R WP 8F is in charge of conducting studies in preparation for Agenda Item 1.4 for WRC 2007. ITU-R WP 8F considers the spectrum implications of the future development of IMT-2000 and provides clarification of spectrum requirements as input to WRC, considering spectrum requirements and recognizing the marketplace and the future development of systems beyond IMT-2000.

B. Capabilities of IMT-Advanced

Recommendation ITU-R M.1645 ([R1]) defines the target peak aggregate data rates for IMT-Advances systems as 1 Gbps for nomadic/local area access and up to 100 Mbps for mobile access. To meet the high aggregate data rate requirements, IMT-Advanced systems will require considerably wider spectrum bandwidths than the current mobile communications systems. Even if the spectral efficiency of IMT-Advanced system will be considerably higher than in the current systems, IMT-Advanced systems will require bandwidths up to 100 MHz to support aggregate data rates up to 1 Gbps. Currently existing bands for IMT-2000 are too narrow and fragmented which does not allow the implementation of 100 MHz carriers. Therefore, the deployment of IMT-Advanced systems with the full envisioned capabilities is not possible on the existing frequency bands.

IMT-Advanced systems will support a wide range of services, which require very high data rates and different quality of service levels. Service classes requiring very high data rates or low latencies will not be covered by existing systems and their enhancements. Market studies predict that there will be demand for services which require up to 50 Mbps or higher per user throughputs in 2010–2020 which cannot be supported by current and evolving systems. For the full deployment of IMT-Advanced systems to answer to meet the needs of future services, new spectrum bands will be needed, which is the topic at WRC 2007.

C. ITU-R Work towards WRC 2007 on Spectrum Requirements

ITU-R WP 8F is in charge of conducting studies in preparation for Agenda Item 1.4 for WRC 2007. As of 2006, ITU-R WP 8F consists of working groups (WG): Future Services & Market Aspects (WG SERV), Developing (WG DEV), Technology (WG TECH) and Spectrum (WG SPEC). ITU-R WP 8F meets three times a year to discuss and decide on the topics. WG SERV has the responsibility to undertake market forecast studies to predict the demand of wireless services in 2010–2020. Especially, WG SERV will give examples of potential services and applications that could be delivered by future developments of IMT-2000 and IMT-Advanced. WG DEV has the responsibility to ensure that the requirements of the developing countries are taken into account. WG TECH has the responsibility to provide technology input to WRC 2007 in terms of potential capabilities and trends in radio access technology. WG SPEC has the responsibility to calculate estimates on the spectrum requirements of IMT-Advanced and evaluate potential candidate bands to full the spectrum requirements.

In preparation for Agenda Item 1.4 for WRC 2007, ITU-R has prepared several Recommendations and Reports:

- Report ITU-R M. 2072: World Mobile Telecommunication Market Forecast ([R3])
- Report ITU-R M. 2074: Radio Aspects for the terrestrial component of IMT-2000 and systems beyond IMT-2000 ([R4])
- Recommendation ITU-R M.1768: Methodology for calculation of spectrum requirements for the future development of the terrestrial component of IMT-2000 and systems beyond IMT 2000 ([P8])

- Tool for estimating the spectrum requirements for future development of IMT-2000 and systems beyond IMT-2000 and the user guide to tool “SPECULATOR” ([P9])
- Report ITU-R M. [IMT.ESTIMATE]: Spectrum requirements for the future development of IMT-2000 and IMT-Advanced ([P10]).

Report ITU-R M. 2072 ([R3]) presents a comprehensive survey of market studies of wireless services in 2010–2020. Report ITU-R M. 2074 ([R4]) presents the grouping of radio access techniques into radio access technique groups and presents the radio related parameters which are used in the spectrum requirement calculations. Recommendation ITU-R M.1768 ([P8]) presents the detailed methodology for calculating the spectrum requirements of future development of IMT-2000 and IMT-Advanced. SPECULATOR ([P9]) is the MS Excel implementation of the spectrum calculation methodology which available at ITU-R WP 8F web site for all ITU members. The spectrum requirements of future development of IMT-2000 and IMT-Advanced have been calculated in Report ITU-R M. [IMT.ESTIMATE] ([P10]) using the methodology described in Rec. ITU-R M.1768 ([P8]), its MS Excel software implementation [P9], input market parameters to the methodology calculations from [R3] and input technology parameters to the methodology from [R4]. ITU-R WP 8F indicated in Report ITU-R M. [IMT.ESTIMATE] that the predicted total spectrum bandwidth requirement for both existing mobile cellular systems, including pre-IMT-2000 and IMT-2000 and its enhancements, and IMT-Advanced for the year 2020 was calculated for both low and high user-demand scenarios to be 1 280 MHz and 1720 MHz, respectively.

D. VTT Participation in ITU-R Work

VTT has participated in ITU-R WP 8F in 2004–2006 in preparation for WRC 2007. VTT representative has attended the ITU-R WP 8F meetings in 2005 as a member of the Finnish delegation. VTT’s emphasis has been on the development of the methodology for calculation of spectrum requirements of future development of IMT-2000 and IMT-Advanced. We have prepared several standardisation contributions to ITU-R WP 8F meetings in 2004–2006 based on the work conducted together with IST-WINNER project. Moreover, we have acted as the official editor of Rec. ITU-R M.1768 ([P8]), which presents the spectrum calculation methodology. We have also participated in the development of software tool [P9] and provided contents and editorial updates to Report ITU-R M. [IMT.ESTIMATE]. Furthermore, VTT contributed to two related publications, [P11] and [P12].

The work at ITU is contribution driven, and different administrations from different parts of the world may have different opinions on the topics. Therefore, it is important to find consensus on the proposals. We have worked in close co-operation with Mobile IT Forum (mITF, Japan's 4G forum) and prepared joint standardisation contributions between Finland and Japan as input to the ITU-R WP 8F work towards WRC 2007.

7.7 Conclusions

The fourth generation (4G) cellular systems emerging during the 2010–2015 have to serve an exponentially increasing number of mobile users and respond to the demand of mobile multimedia services with high-data-rate up to 100–150 Mbits/s. This chapter summarized the power control research for cellular and wireless local area networks, fundamental stochastic features of network simulations for the 4G systems development, and depicts the path for the development and standardization of appropriate technologies for the future cellular systems.

Publications

[P1] T. Frantti, "Multiphase transfer of control signal for adaptive power control in cdma systems." Control Engineering Practice, Elsevier Science, Vol. 14/5, pp. 489–501, 2005.

[P2] T. Frantti, "Power control method and telecommunications system." Accepted Patent, 020610EPO Europe, 20020610US USA, 20020610JP Japan, 2005.

[P3] T. Frantti, P. Mähönen, "Adaptive fuzzy power control for wcdma mobile radio systems." Control Engineering Practice, Vol. 9, pp. 947–959, 2001.

[P4] J. Hiltunen, "Channel Reservation in a Power Controlled Wireless Local Area Network." Master's thesis, University of Oulu, Department of Electrical Engineering, Oulu, Finland, 2005.

[P5] T. Frantti, "Reliable simulation of communication networks." WSEAS Transactions on Communications, Issue 5, Vol. 5, pp. 758–765, 2006.

[P6] M. Taramaa, "Performance Optimization of Ad Hoc Network." Master's thesis, University of Oulu, Department of Electrical Engineering, Oulu, Finland, 2005.

[P7] T. Frantti, M. Taramaa, "Fundamental features of ad hoc networks' simulations." WSEAS Transactions on Communications, Accepted for Publication, pp. 758–765, 2006.

[P8] Draft new Recommendation ITU-R M.1768 "Methodology for calculation of spectrum requirements for the future development of the terrestrial component of IMT-2000 and systems beyond IMT-2000." Editor M. Matinmikko, 2006.

[P9] ITU-R WP 8F Document 8F/900, "Tool for estimating the spectrum requirements for future development of IMT-2000 and systems beyond IMT-2000 and the user guide to tool." 2006.

[P10] Draft new Report ITU-R M. [IMT.ESTIMATE] "Spectrum requirements for the future development of IMT-2000 and IMT-Advanced." 2006.

[P11] T. Irnich, M. Matinmikko, J. Huschke, J. Ojala, P. Ojanen, B. Walke, S. Thilakawardana, "Methodology for calculating the spectrum requirements for the future development of imt-2000 and systems beyond imt-2000." In: Wireless World Re search Forum Meeting 14, San Diego, California, 2005.

[P12] T. Irnich, M. Matinmikko, B. Walke, "Spectrum requirements for imt-2000 using the new itu spectrum requirement estimation methodology." In: Wireless World Research Forum Meeting 14, San Diego, California, 2005.

References

[R1] Recommendation ITU-R M. 1645 “Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000.” 2003.

[R2] Recommendation ITU-R M.1390 “Methodology for the calculation of IMT-2000 terrestrial spectrum requirements.” 1999.

[R3] Draft new Report ITU-R M.2072 “World mobile telecommunication market forecast.” 2005.

[R4] Draft new Report ITU-R M.2074 “Radio Aspects for the terrestrial component of IMT-21000 and systems beyond IMT-2000.” 2005.

8. Wireless World Initiative New Radio II (WINNER II)

Jussi Roivainen, Marja Matinmikko, Aarne Mämmelä, Yan Zhang,
Chi Hua Geng, Miia Mustonen, Ilkka Saarinen and Jahangir Sarker

8.1 Project Goals

The goal of WINNER is to develop a single ubiquitous radio access system adaptable to a comprehensive range of mobile communication scenarios from short range to wide area. This will be based on a single radio access technology with enhanced capabilities compared to existing systems or their evolutions.

WINNER II is a continuation of the WINNER I project, which developed the overall system concept. WINNER II will develop and optimise this concept towards a detailed system definition. All investigations will take place within the context of a system view to enable a focused development of a system rather than individual components. In addition limited trials will be performed in order to assess some key elements of the WINNER II system.

The radio interface will support the challenging requirements of systems beyond 3G. It will be scalable in terms of carrier bandwidth and carrier frequency range and it will allow deployment in newly identified and “re-farmed” frequency bands. The system will support a wide range of usage and radio environments providing a significant improvement in performance and Quality of Service. The resulting system specification will meet future market demands and will provide optimum user experience.

The radio interface will optimise the use of spectral resources, e.g. through the exploitation of actual channel conditions and multiple antenna technology. New networking topologies (e.g. relaying) will support cost-effective deployments. Support of advanced resource management and handover will ease the deployment of the WINNER system enabling seamless service provision and global roaming.

The project will also contribute to the global research, regulatory and standardisation activities. The WINNER approach targets a globally harmonised system. The project objectives are shared by a strong Consortium of major players in the mobile and wireless communication industry including manufacturers, network operators, R&D centres and the academic domain.

VTT participated in WINNER I to two Work Packages, namely WP6 Spectrum and Coexistence and WP7 System Engineering.

The main objectives for VTT in WP6 Spectrum and Coexistence are

- To develop a methodology for estimating the spectrum requirements for "further developments of IMT-2000 and systems beyond IMT-2000". suitable for the ITU in it's preparations towards the WRC-2007
- The developed methodology will be used to prepare an estimation of the spectrum requirements for (i) the systems beyond IMT-2000 as well as for (ii) the WINNER system concept.

The main objective for VTT (together with other WP participants) in WP7 System Engineering were to enable and ensure a realisation of the listed WINNER project objectives. To that effect, WP7 encompasses the technical and supporting components and synthesises their outputs. The following basic goals exist:

Technical guidance of the IP:

- To identify the system requirements, which are used to drive the technical scope of the overall project
- To develop and provide assessment focus to the other Work Packages
- To steer the technical direction of Work Packages to achieve system consistency and to fulfil new radio system requirements
- To provide technical facilitation across Work Packages.

Synthesis activities in which all WINNER I activities are focused:

- To receive scenarios, radio access and radio network technologies, spectrum and coexistence issues, identified in other Work Packages, as an input

- To base on the inputs and to compile basic system concept envisaged as an appropriate mapping between scenarios and technologies
- To compile an initial complexity evaluation based on inputs from Work Packages.

VTT participates in WINNER II to two work packages: WP5 Spectrum and WP6 System concept. WINNER Phase II is focused on the detailed system design, optimisation, validation (through limited trials) and preparation of further trials in Phase III. This translates into the following WINNER II objectives, which are related to the work where VTT participates:

- To design and optimise the new ubiquitous radio access system, whose parameters can be scaled or adapted to the requirements of a comprehensive range of mobile communication scenarios. From a coverage area point of view wide-area, metropolitan and short-range scenarios have to be supported. The radio access system should be capable of supporting variable bit rates, with peak data rates of up to approximately 100 Mbps for medium to long-range heavy traffic areas with high mobility, and up to approximately 1 Gbps for hot spots and short-range scenarios.
- To contribute to the international standardisation and regulatory process – in particular to the development of the necessary reports and recommendations in ITU-R in the preparatory phase of WRC 2007 – and where appropriate to other international bodies, where WINNER II can provide technical input. One example might be the 3G evolution study item in 3GPP on special topics depending on the detailed workplan in 3GPP.
- To perform limited in order to proof the concept of basic functionalities of the WINNER II system.
- To prepare the trial campaign by the selection of the hardware and software platform of the trial system and preparatory activities towards the application of a frequency test license.
- To disseminate results via international conferences, reputable journals and the organisation of workshops as part of a global harmonisation process.

8.2 Project Work

8.2.1 WINNER System Capabilities and Spectrum Requirements

The vision of a ubiquitous radio system concept in WINNER is one of providing wireless access for a wide range of services and applications across all environments, from local area to wide-area, with one single adaptive system concept for all envisaged radio environments. It will efficiently adapt to multiple scenarios by using different modes of a common technology basis. This is built on the recognition that developing separate systems for different purposes (cellular, WLAN, short-range access etc.) will no longer be optimal solution in the future converged wireless world where overall optimisation can maximise the performance and ensure full functionality for wide range of services over the full range of environments.

The WINNER system concept will comprise of optimised combination of the best available components, based on an analysis of the most promising solutions and concepts available or proposed within the research community. The initial development of new technologies and their combination in the system concept will be further advanced with respect to the project goals towards future system realisation.

Compared to current and evolving mobile and wireless systems, the WINNER system concept aims to provide significant improvements in peak data rate, latency, mobile speed, spectrum efficiency, coverage, cost per bit and supported environments taking into account specified Quality-of-Service requirements.

Future growth of mobile and wireless communications is expected mainly from data-oriented services and applications. For future business scenarios aggregate throughputs between 100 Mbps and 1 Gbps are foreseen as suitable minimum requirement in Recommendation ITU-R M.1645 ([R1]) for the new mobile access capabilities and new nomadic/local area wireless access capabilities. Service classes requiring very high data rates or low latencies will not covered by existing systems and their enhancements. The enhanced capabilities of the WINNER system will offer features which are not possible in current cellular systems like full coverage of higher sustainable data rates together with good QoS performance and full mobility support, e.g. data rates and performance comparable to wire line connections of today. The system will also offer a significant improvement in spectral efficiency compared to current cellular systems.

Requirements from Recommendation ITU-R M.1645 ([R1]) imply the need for a new flexible radio access system with much higher peak data rate and other performance, requiring higher carrier bandwidth in order to provide suitable user satisfaction. New and more efficient radio access technologies will be needed for supporting ubiquitous communication at a better cost-benefit-ratio compared to existing systems.

The WINNER radio interface has been designed as a packet-oriented always-best concept. It defines a scalable and flexible interface based on adaptive system modes tailored to particular situations such as the radio environment, the usage scenario, the economic model, etc. Parameterizing of the modes allow for additional flexibility. According to current working assumptions, key components of the radio interface include:

- Ability to handle a wide range of bandwidths, from 1.25 MHz to 100 MHz, at carrier frequencies up to 6 GHz.
- Duplex choice – it supports both paired spectrum via frequency-division duplex (FDD, supporting half-duplex terminals) and unpaired spectrum via time-division duplex (TDD).
- Generalised Multicarrier (GMC), which enables flexible switching between different forms of multi-carrier and single carrier modulation depending on the specific needs, at minor extra implementation complexity. GMC includes standard cyclic-prefix orthogonal frequency division modulation (CP-OFDM) and single carrier transmission as special cases.
- Advanced multi-antenna systems and spatial processing as an integral part of the concept, including components such as MIMO and beam forming and several other possibilities within a common generic framework.
- Relay-enhanced cells as an integrated part, extending coverage and capacity.
- A multi-mode protocol architecture.

In addition, adaptivity on all time scales and very short transmission delays are important features of the concept.

The WINNER system has the flexibility to handle a wide range of bandwidths, in both paired and unpaired bands, and to operate on different modes, which allow adaptation to a wide range of situations and environments. Generalised multicarrier (GMC) transmission scheme enables flexible switching between different forms of multicarrier and serial modulation allowing efficient use of any available bandwidth.

The WINNER system also includes multihop capability and advanced multi-antenna concept and spatial processing to provide good coverage and capacity. Spectrum sharing and FSU are important aspects of the WINNER system concept providing flexibility, efficient use of the spectrum and ability to co-exist with other systems.

The frequency spectrum is the fundamental resource of mobile and wireless communication. Sufficient spectrum is the prerequisite for successful deployment of new systems and establishment of associated new business areas. Due to physical limitations, advances in radio technology can not provide sufficient spectrum efficiency to compensate the future high data rate requirements so that the existing bands would be sufficient. The existing bands are too narrow and too fragmented to realize the substantial potential of this new radio technology. Therefore, it is important that World Radiocommunication Conference WRC 2007 considers the frequency related matters of 4G systems and possibly identifies sufficient spectrum to allow 100 MHz wide channels. This would allow proper deployment of the future mobile broadband technologies, such as WINNER, with the capabilities envisioned in Recommendation ITU-R M.1645 ([R1]). The preference from technical perspective thus is for a single continuous frequency assignment for systems beyond IMT-2000. This would help to avoid inter-system guard bands and would also be beneficial for reducing implementation complexity on both infrastructure and terminal side. The WINNER system capabilities and spectrum requirements are further discussed in [P1].

ITU has started the standardization process for IMT-Advanced. Contributions are expected from several countries about the technical minimum requirements of the IMT-Advanced system related to radio, spectrum and services. WINNER project is also preparing its contributions to the standardization process and VTT is involved in this activity. At the beginning of the year 2008 the agreed technical minimum requirements for the IMT-Advanced systems will be gathered into the circular letter which will be sent out to as an invitation from the Radiocommunication Bureau (BR) to the member states, sector members and to the external organizations to propose candidate radio interface technologies for IMT-Advanced. The candidate radio interface technologies will be evaluated and the ones that fulfil the technical minimum requirements will be accepted as an IMT-Advanced, thus, 4G systems. WINNER system is aiming to one of the accepted ones.

8.2.2 Low Density Parity Check (LDPC) Codec Implementation

VTT's main contributions in the WINNER2 project's work package 6 are making of a specification for the low-density parity-check (LDPC) codec blocks. The LDPC should not be only suitable for the trial system demonstration in T12, but also flexible for future extensions in other projects. Another part of the work package 6 is to implement the specified functionalities of the LDPC codec according to the hardware constraints. Currently the encoder is being implemented and decoder implementation work is planned for spring 2007.

LDPC codes were first introduced in 1963 by Gallager ([R2]). The interest to LDPC codes has been limited until recently because of LDPC codes' high storage requirements and complexity. During 1990's LDPC codes were independently rediscovered by both MacKay ([R3]) and Wiberg ([R4]). LDPC codes can be the next step after turbo codes now when technology has evolved to a level that can handle the resource requirements of LDPC codes. LDPC codes have an asymptotically better performance than turbo codes and provide a wide range of tradeoffs between performance and decoding complexity. Besides the technology evolution, the theoretical improvement is pushing LDPC to applications. LDPC is appearing in standards such as IEEE 802.16e (draft) and DVB-S (digital video broadcasting -satellite). Consequently, LDPC codes implementation is becoming a practical topic in industry R&D.

The work was started by performing a thorough literature review of the LDPC algorithm. This includes the study of previous WINNER deliverables, finding out the state-of-the-art for this advanced channel coding algorithm, feasibility studies concerning the implementation workload, and a collection of comparable designs. Frequent contacts with other partners in work package 2 of the WINNER II project was an important part of the specification work since quite a lot research work has been done in work package 2 during the WINNER 1 project. Besides the WINNER I research we needed the newest information from these partners in order to finish the specification that is found on literature. Also contacts with other partners inside are important to find some specification parameters and interface definitions for the work package 6 trial demonstrator.

The detailed studies on the LDPC implementation architectures identify the bottlenecks of the implementation work. The LDPC is implemented using standard VHDL language, simulation and synthesis processes. The resulting code is integrated and tested to ensure synthesizable IP blocks for LDPC encoder and decoder which can be integrated quite easily to the rest of the trial system.

8.3 Project results

8.3.1 Methodology for Calculation of Spectrum Requirements of IMT-Advanced

In the past, estimation of spectrum requirements of wireless applications has been considered as a framework focusing on a single system and market scenario. With the advent of a convergence of mobile and fixed telecommunication and multi-network environments as well as supporting attributes like seamless interworking between different complementary access systems, as described in Recommendation ITU-R M.1645, application of such a simple approach is no longer suitable. For the estimation of frequency requirements, new models have to be developed and applied.

WINNER system concept was defined and complexity estimation done in WINNER I project. Terminal power consumption is modelled with a spreadsheet model. Also transmitted energy as basic system resource ([P2]) was studied.

According to Agenda Item 1.4, the next WRC 2007 will consider frequency related matters for the future development of IMT-2000 and systems beyond

IMT-2000. As preparation for the conference, ITU has developed a methodology ([P3]) to calculate the spectrum requirements of future development of IMT-2000 and systems beyond IMT-2000 (i.e. IMT-Advanced or 4G systems). The ITU methodology is to a large extent based on the results from the WINNER project.

The starting point for all spectrum considerations concerning IMT-2000 and systems beyond IMT-2000 are the market expectations for wireless communications services between 2010 and 2020. The methodology takes a technology-neutral approach in its technical studies of radio access techniques and uses the classification of radio access technique groups (RATGs) defined in Report ITU-R M.[IMT.RAD_ASPECTS] ([R5]). The spectrum calculation methodology requires technical parameters to characterize the different RATGs as input to the spectrum calculations. By the RATG approach, the technical consideration for spectrum estimation can easily be conducted without referring to the detailed specification of radio interfaces both of existing and future mobile systems.

The methodology accommodates a complex mixture of services from ITU-R market studies [R6] with service categories having different traffic volumes and QoS constraints. The methodology takes into account the time varying and regionally varying nature of traffic. The methodology applies a technology neutral approach to handle emerging as well as established systems using radio access technique group (RATG) approach with a limited set of radio parameters. Four RATGs considered cover all relevant radio access technologies. The methodology distributes traffic to different RATGs and radio environments (i.e. cell layers) using technical and market related information. The methodology considers all wireless access technique groups but calculates the spectrum requirements only for RATGs 1 and 2, which correspond to pre-IMT, IMT-2000 and its enhancements, and IMT-Advanced, respectively. For the traffic distributed to RATGs 1 and 2 the methodology transforms the traffic volumes from market studies into capacity requirements using separate algorithms for packet-switched and circuit switched service categories and takes into account the gain in multiplexing packet services with different QoS characteristics. The methodology transforms capacity requirements into spectrum requirements using spectral efficiency values. The methodology considers practical network deployments to adjust the spectrum requirements and calculates the aggregate spectrum requirements for further development of IMT-2000 and systems beyond IMT-2000.

Most parts of the spectrum calculation methodology in Recommendation ITU-R M.1768 ([P3]) were developed in the WINNER project. We have demonstrated the methodology developed in WINNER project to large audiences at 14th IST Summit in Dresden in June 2005 and Wireless World Research Forum Meeting 15 in Paris in December 2005. The methodology has been implemented in MS Excel in the WINNER project. The developed software tool [P4] is now available for all ITU members and ITU has used it in estimating the spectrum requirements for WRC 2007 in [R7].

8.3.2 Systems Engineering Results

In addition to spectrum, energy is a basic resource in all digital transmission links ([P2]). Energy is used in signal processing and transmission. Physically, radio channels correspond to passive circuits, which cannot generate any new energy. In fact most of the transmitted energy is lost in the channel. Radio waves follow the energy conservation law, but the transmitted energy is distributed in all directions and is thus mostly unavailable to the receiver, or the energy changed into other forms in the transmission medium, for example into heat. Only a small part of the energy is actually received and the signal is seen as highly attenuated. Quite easily the attenuation can be as high as 80–100 dB.

We can use two alternative approaches for performance measurements in terms of energy. Either the average transmitted or received energy per bit is used, both usually normalized by the receiver noise spectral density. This leads to the average transmitted or received signal-to-noise ratio (SNR) per bit, respectively. However, it is the transmitted energy which is the basic system resource since in mobile terminals the transmitted energy is taken from the battery with a finite capacity. The received energy is only a small part of the transmitted energy. The ratio of the received energy to the transmitted energy is the energy gain of the channel. For normalization purposes the representative energy gain is defined as the average energy gain for a signal that is uniformly distributed in time, frequency and space. The average energy gain of a channel depends on how well the transmitted signal is "matched" to the channel. In general, water filling predicted by the information theory gives a better average energy gain than truncated channel inversion. Thus contrary to the common belief the channel is not simply a scaling factor for the energy and it is crucial which energy is used in comparisons.

For convenience, we defined the transmitted SNR referred to the receiver as the product of the transmitted SNR and the representative energy gain. We derived an explicit relationship between the transmitted and received SNR's using the covariance concept. Limitations of the use of different SNR definitions were summarized. We gave examples which showed that the use of received SNR give completely misleading results when transmitter power control is used although many authors have used the received SNR in recent archive journals.

An additional problem is the proper normalization of the fading channel model [P5]. The major approaches for normalization include setting of either the average representative energy gain or the peak energy gain to unity. The peak energy gain of many fading models including Rayleigh fading is infinite, which is obviously impossible in a passive system where the peak energy gain should be less than or equal to unity. Our aim was to show that it is due to the average normalization that quite unexpectedly the performance in a fading channel seems to be better than in a nonfading channel at low transmitted SNRs. With peak normalization this does not happen. One benefit from the peak normalization is that the minimum required transmitted energy can be directly derived from the Shannon limit. Our results can be used in fair comparisons of adaptive transmitters and also in comparing the energy efficiency of wireless networks.

8.3.3 LDPC Codec Implementation Results

The first result is a complete specification for the LDPC codec implementation. The specification covers the code design perspective, as well as the algorithmic decisions and the base model matrices definitions for different code rates. Table 1 shows the parameters and decisions that have been chosen for the implementations. Comparisons between the Block LDPC, turbo code and convolutional code have been conducted in work package 2, which lead to many instructive conclusions. Thus most of the decisions in Table 1 in the code design and algorithm parts are following the results from work package 2. However, in order to have an opportunity to reuse this LDPC block in the near future for other VTT's projects, we decided to set the code rate and code length real-time adjustable. Architectures that have been chosen for the encoder and decoder are considered to fulfill a reasonable performance with technically feasible resources in the trial system. Finally, the decisions for the hardware implementation constraints are mostly made by task 12 leader in order to have an easy integration of the LDPC IPs and the rest trial system.

Table 1. Parameters and decisions for LDPC implementation.

Abstraction Level	Parameters	Current Decision
Code Design	Structure	Block-LDPC
	Code rate	Real-time selectable between 1/2, 2/3 and 3/4
	Code length	Real-time selectable between 576, 1152, 1728, 2304 and 4608 bits
	Base model matrices	Defined in WP2 in Winner (Note: this is different from the matrices defined in D2.10)
	Expansion rules	Modulo
Algorithm	Encoding algorithm	Modified version of the Richardson paper, following the pipeline structure. (Note: reference [R8], [R9])
	Decoding algorithm	Corrected Min-Sum, scaling factors
	Scheduling algorithm	Horizontal shuffling
Architecture	Encoder	pipelined
	Decoder	partial-parallel
Implementation	Device	Xilinx Virtex 2 6000
	Data rate	10–40 Mbps
	Quantization	uniform
	Internal data width	4 to 5 bits
	Input to the decoder	8-bit soft symbol
	Output from the decoder	hard-decision bit

The second result is a literature review for the LDPC codes as well as the encoder implementation work that will be reported later on a WINNER project deliverable. The architecture of the encoder follows the pipelined structure that is described by Richardson and Urbanke in 2001 ([R8]). The encoder receives the information block $\mathbf{s} = (s_0, \ldots, s_{k\text{-}1})$ and uses the expanded matrix **H** to determine the parity-check bits. The expanded matrix **H** is determined from the base model matrix defined in work package 2. For efficient encoding, H is divided into the form

$$H = \begin{bmatrix} A & B & T \\ C & D & E \end{bmatrix} \tag{1}$$

Due to the nice cyclic structure of the Block LDPC, the pipeline structure discussed in [R8] and [P1] could be further simplified as the one shown in the above part of Figure 1. Basically each gray rectangle in this figure means a matrix multiplication. L letters A, B, C, E, T refer to the sub-matrices defined in (1). The symbol ET_I refers to the inversion matrix of E*T and T_I refers to the inversion matrix of T. These matrices could all be derived from the expanded matrix **H.** The lower part of Figure 1 shows a more detailed view of the above modified pipeline structure. Due to the cyclic property of the Block LDPC, instead of multipliers the matrix multiplication can be realized by an XOR operation and shifting. Some small tests with Matlab and VHDL have been conducted in order to verify this simplified approach.

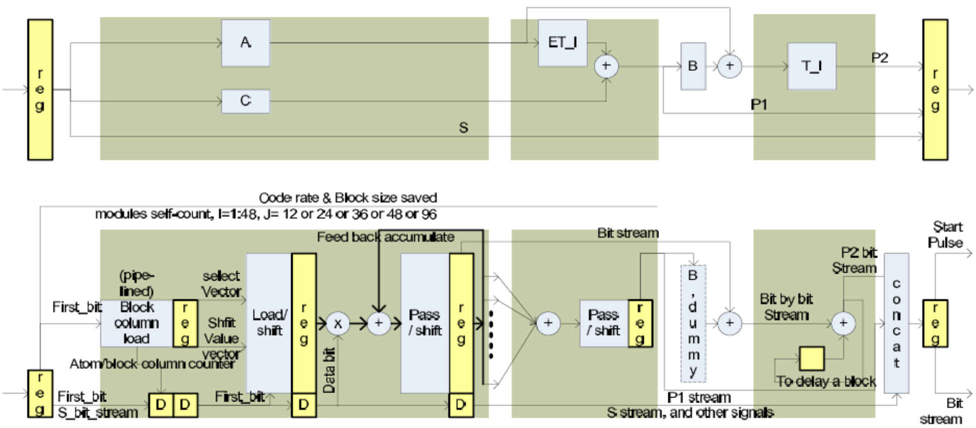


Figure 1. The encoder block diagrams at two different levels.

8.4 Conclusion

The WINNER vision is a ubiquitous radio system concept that provides wireless access for a wide range of services and applications across all environments, from local area to wide-area, with one single adaptive system concept for all envisaged radio environments. It will efficiently adapt to multiple scenarios by using different modes but use a common technology basis. The development of separate systems for different purposes (cellular, WLAN, short-range access etc.) will no longer be an optimal solution in the future converged wireless world, because overall optimisation can maximise the performance and ensure full functionality for wide range of services over the full range of environments.

Future growth of mobile and wireless communications is expected mainly from data-oriented services and applications like video streaming. For future business scenarios aggregate throughputs between 100 Mbps and 1 Gbps are foreseen as suitable minimum requirement in Recommendation ITU-R M.1645 for the new mobile access and nomadic/local area wireless access capabilities. The existing systems and their anticipated enhancements do not cover the service classes requiring very high data rates or low. The enhanced capabilities of the WINNER system will offer features which are not possible in current cellular systems, for example a full coverage of higher sustainable data rates together with good QoS performance and full mobility support, that is, data rates and performance comparable to wire line connections of today. The system will also offer a significant improvement in spectral efficiency compared to current cellular systems.

VTT's participation to WINNER I and WINNER II has resulted input to spectrum work and that work continues with WINNER II project to end of year 2007. The work serving ITU spectrum requirements is a significant input for future spectrum efficiency and network co-operation possibilities. The VTT's effort, together with other WINNER partners, has produced several results, for example the updates of the spectrum calculator tool SPECULATOR. We have also contributed actively to ECC PT1 and ITU-R recommendations.

VTT's contribution in the WINNER2 project's work package 6 is the definition of specifications for the low-density parity-check (LDPC) codec blocks. The specified LDPC is not only suitable for the trial system demonstration in WINNER, but also flexible for future extension. Another part of VTT's work is to implement the specified functionalities of the LDPC codec according to the hardware constraints. Currently the encoder is being implemented using the pipelined structure. The decoder implementation work is planned for spring 2007.

Publications

[P1] W. Mohr, J. Axnäs, A. Gomes, J. Lara, C. Legutko, M. Matinmikko, P. Ojanen, S. Pfletshinger, "WINNER System Concept: Capabilities and Spectrum Usage." Wireless World Research Forum Meeting 16, Shanghai, China, 26–28 April 2006.

[P2] A. Mämmelä, I. Saarinen, D. P. Taylor, "Transmitted energy as a basic system resource." Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE, Vol. 6, 28 Nov. – 2 Dec. 2005. 5 p.

[P3] "Recommendation ITU-R M.1768 Methodology for calculation of spectrum requirements for the future development of the terrestrial component of IMT-2000 and systems beyond IMT-2000." Editor M. Matinmikko, 2006.

[P4] "Tool for estimating the spectrum requirements for future development of IMT-2000 and systems beyond IMT-2000 and the user guide to tool." Available for ITU members at ITU-R WP 8F web site: http://www.itu.int/ITU-R/study-groups/index.asp?link=rwp8f&lang=en (Contact info J. Ojala & M. Matinmikko)

[P5] A. Mämmelä, M. Höyhtyä, D. P. Taylor, "Normalization of a fading channel." Accepted for publication in Proc. GLOBECOM'06, IEEE Global Telecommunication Conference.

References

[R1] "Recommendation ITU-R M.1645 Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000." 2003.

[R2] R. G. Gallager, "Low-Density Parity-Check Codes". Cambridge, MA: MIT Press, 1963.

[R3] D. J. C. MacKay, R. M. Neal, “Near Shannon limit performance of low density parity check codes.” Electron. Lett., Vol. 32, Aug. 1996, pp. 1645–1646.

[R4] N. Wiberg, “Codes and decoding on general graphs.” Dissertation No. 440, Dep. elect. Eng. Linköping Univ., Linköping, Sweden, 1996.

[R5] “Draft new Report ITU-R.[IMT.RAD_ASPECTS] Radio Aspects for the terrestrial component of IMT-21000 and systems beyond IMT-2000.” November 2005.

[R6] “Draft new Report ITU-R.[IMT.MARKET] World mobile telecommunication market forecast.” November 2005.

[R7] “Report ITU-R M.[IMT.ESTIMATE]: Spectrum requirements for the future development of IMT-2000 and IMT-Advanced.”

[R8] T. J. Richardson, R. L. Urbanke, “Efficient encoding of low-density parity-check codes.” IEEE Trans. On Info. Theory, Vol. 47, N. 2, Feb. 2001, pp. 638–656.

[R9] S. Myung, K. Yang, J. Kim, “Quasi-Cyclic LDPC codes for fast encoding.” IEEE Trans. On Info. Theory, Vol. 51, No. 8, Aug. 2005, pp. 2894–2901.

9. Measurement Tool for Mobile Geographic Information System (Mobilised MGIS)

Ali Lattunen, Paul Kemppi, Juuso Pajunen, Yanli Li and Seppo Horsmanheimo

9.1 Introduction

In the Mobilised MGIS project, a portable software tool, Wireless Network Measurement Kit (WiNK), capable of collecting terminal level network information from an attached mobile device, was designed and implemented. The collected data samples are combined with reference locations given by GPS receiver and transferred to the server over either wireless or wired TCP/IP connection.

Well known tools exist in the market for collecting signal information from cellular networks, such as Nemo Outdoor [R2] and Tems Investigation [R3]. They are rather expensive and also too complicated for many purposes. There are also tools for collecting data from WLAN networks, such as Network Stumbler [R4]. However, by the time we started developing the measurement tool, the existing tools did not support collecting measurement samples simultaneously from GSM and WLAN networks. Thus, the developed software, as a lightweight measurement tool yet with sufficient measurement functionalities, and with illustrative visualisation of measured data, will definitely benefit researchers and network administrators in the fields of network planning and mobile positioning.

In this report, we first introduce the architecture and functionality of the MGIS software in section 9.2. Then, in section 9.3, we describe system evaluation through a positioning trial conducted in Otaniemi, Espoo (Finland) and present shortly the utilised positioning method called Database Correlation Method (DCM). In Section 9.4, we show the results of the positioning trial and prove that the accuracy of the Database Correlation Method (DCM) can be enhanced by combining signal information from GSM and WLAN networks collected with the implemented measurement tool. Finally, we wrap this chapter up with conclusions in section 9.5.

9.2 MGIS Software

The overall MGIS system architecture consists of server components, Location Server and MGIS Web Service (MGIS-WS), and a rich client WiNK. However, the WiNK client is independent enough to fully work without connection as well. The MGIS architecture is shown in Figure 1.

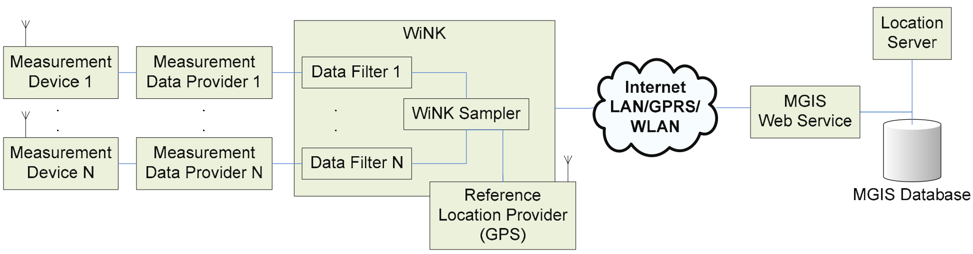


Figure 1. MGIS system architecture.

9.2.1 WiNK

WiNK is a multi-functional tool for network professionals out in the field. It is comprised merely of software utilising publicly available commercial hardware. Thus, it is lightweight, easily deployable and inexpensive to use. Since it is network operator independent as well, it is easily customisable to be used for collecting terminal level measurement data from different cellular networks. At the same time, WiNK is an effective analysis tool for verifying the results of novel coverage estimation and mobile positioning methods. A screenshot of WiNK user interface is shown in Figure 2.

The WiNK tool acts as a client and takes advantage of the MGIS server both as a central storage of measurement data and as a provider of location estimations calculated according to those data. When acting as a measurement tool, WiNK periodically gathers data samples from the measurement data providers that may be interfacing any kinds of applicable wireless network terminals. The only practical limitation for the terminal is that it must comprise a programmatically accessible interface for collecting network level data. The measurement data provider modules have been segregated from WiNK as their own libraries, so the measured network types can be easily changed by replacing them.

1) *Measurement data providers*: GSM and WLAN signal strengths are collected by using separate data provider libraries. The current GSM data provider implementation supports NetMonitor [R5] enabled mobile phones complying with the packet based Fbus V2 command protocol [R6]. For this reason, not all cellular network terminals are suitable for being used as WiNK system measurement devices straight away. Another restriction related to the measurement equipment is the moderately low data sampling frequency due to the limited processing capabilities of mobile phones. This may result in position-wise sparse data when measuring in a fast moving vehicle. The WLAN data provider obtains the received signal strengths of WLAN APs through Windows NDIS 5.1-interface [R7].

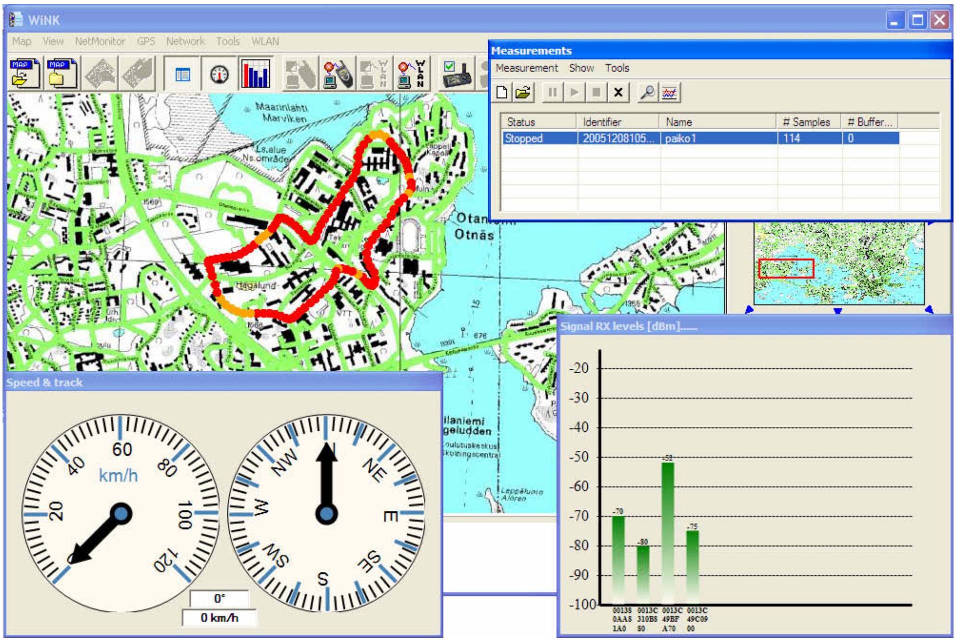


Figure 2. WiNK user interface.

2) *Data collection:* The sampler module inside WiNK listens to the events originated by measurement data providers and carrying measurement samples, and periodically triggers these values to the storage. Each signal fingerprint sample is coupled with both a timestamp and navigation information (reference 3-D location, speed and heading) provided by the Reference navigation info provider, which is presently based on GPS

navigator in our system. When collecting measurement data, filtering methods can be applied to the measurement data in order to gather data about either selected or specific types of carriers. Several types of measurements can be active and running simultaneously.

3) *Geographic visualisation:* In regard to geographic visualisation, the WiNK software leans on NETCharts.VTT, our digital map library designed to facilitate and automate raster map handling, geographic visualisation and map based navigation. The NETCharts.VTT library automates importing and calibrating of raster map images; it internally maintains a both geographically and scale-wisely ordered map index, which allows both scale-wise and position-wise navigation between the loaded maps. NETCharts.VTT also provides means for complementing map information by placing graphical annotation objects, such as texts, shapes and symbols, on successive transparent layers to be displayed on top of the map images.

4) *Visualisation of results:* When WiNK is used as a verification and visualisation tool for estimated positioning results, graphical comparison of estimated locations and corresponding reference locations are displayed on the appropriate digital map. The estimations are determined for the same positions as where the measurement samples have been collected, which ensures the comparison validity. Some simple statistical data from the comparison results are available as well.

5) *External interfaces:* As WiNK is mainly intended for measurement purposes, it relies on commercial products in advanced presentation and analysis capabilities. An interface to ArcGIS [R8] software allows exporting measurement data in text formats that can be imported to ArcMap [R11] for further processing. The maps presented later in this chapter were created with ArcGIS.

9.2.2 MGIS Web Service and Location Server

MGIS-WS acts as a server for the WiNK client. WiNK currently uses MGIS-WS for three purposes: for storing fingerprint data to the database, for retrieving those data from the database and for location calculation. The fingerprint data can be stored in two ways: in a batch mode, one measurement at a time, or each sample can be processed individually and stored to the database in real-time.

MGIS-WS provides the location calculation service using the Location Server component. The location methods supported by LS can be queried through MGIS-WS. The desired methods can then be invoked, either separately for one sample or for all the samples of one measurement.

9.3 System Evaluation

To be able to evaluate the functionality of the MGIS software and the validity of the data collected by it, a positioning trial was carried out in Otaniemi, Espoo (Finland). The trial consists of coverage area measurements, test route measurements to be used in positioning and deployment of the DCM algorithm.

9.3.1 Measurements

The measurements were carried out in Otaniemi, which is a suburban area with relatively open signal propagation environment. It mainly comprises low-rise buildings with wide spacing as shown in Figure 3.

In the Otaniemi area, TeliaSonera's [R9] GSM network consists mainly of micro and small cells, but due to the open sea around the cape of Otaniemi, some GSM cells are also heard over the bay from the north. The density of WLAN APs in Otaniemi is quite high because of the campus networks of the Helsinki University of Technology (TKK). WLAN APs reside generally in entrance halls, corridors and auditoriums, but their coverage areas stretch out through the windows.

Figure 3. Three-dimensional image of the Otaniemi area based on data provided by TerraSolid.

Before the actual positioning trial, three WLAN cards were tested with WiNK. The performance of the cards varied significantly. All the cards performed well when being stationary or moving slowly. However, as movement speed increased, causing detected WLAN APs to change more frequently, some cards could not keep up. The tested cards and their perceived performance are listed in Table 1.

Table 1. Wireless card performance.

Card	Driver version	Comments
Intel Pro Wireless 2200 BG	9.0.3.9	Incorrectly returns WLAN APs faded out of sight.
D-Link DWL-G650+	2.02	Does not return all WLAN APs.
Linksys WPC54G	3.100.64.0	OK.

After finding the properly functioning WLAN card, a coverage measurement was carried out by using WiNK to collect the reference fingerprints for the DCM algorithm. It took approximately 2.5 hours by car to cover all the streets in the area.

In addition, the road around the Laajalahti bay (north of Otaniemi) was covered. The reference fingerprints within the Otaniemi area are depicted in Figure 4.

Altogether, 188 different GSM cells and 357 WLAN APs were detected in the coverage measurement area. The average number of detectable GSM cells per fingerprint was 6.1 and the measured average signal strength was -74.8 dBm with standard deviation of 10.8 dBm. WLAN APs were understandably detected only in the vicinity of buildings (see Figure 4). The average signal strength for the detected WLAN APs was -85.1 dBm with standard deviation of 5.6 dBm.

The reference fingerprints were averaged with 5-meter grid spacing in order to reduce the amount of fingerprints used in correlation calculation and thus, to improve the efficiency of DCM. After averaging, the total amount of reference fingerprints was reduced from 2369 to 1795.

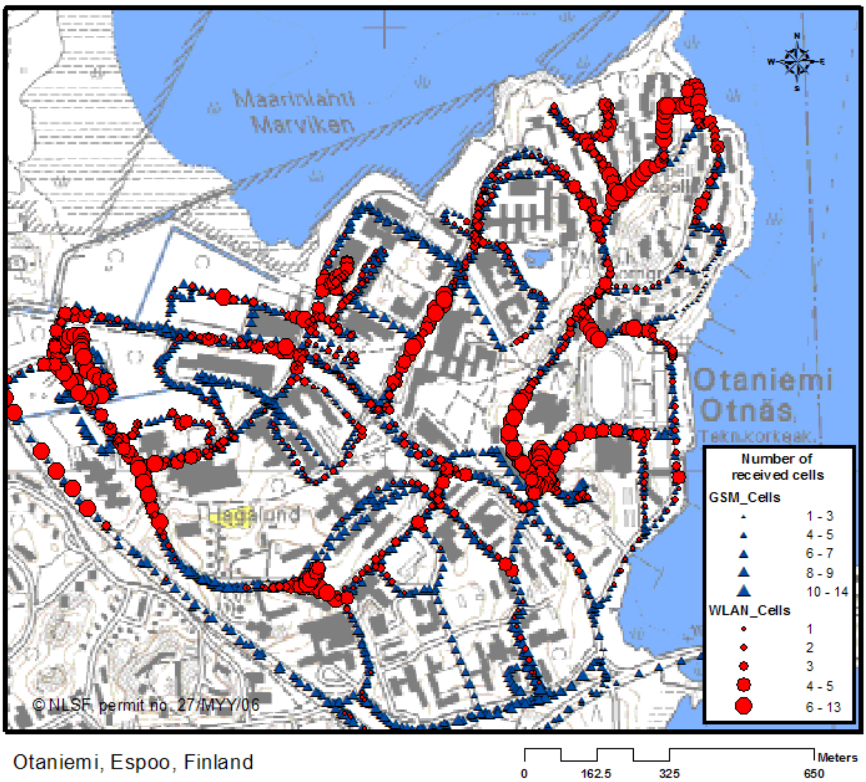


Figure 4. The distribution of received GSM cells and WLAN APs for the reference fingerprints in the Otaniemi area.

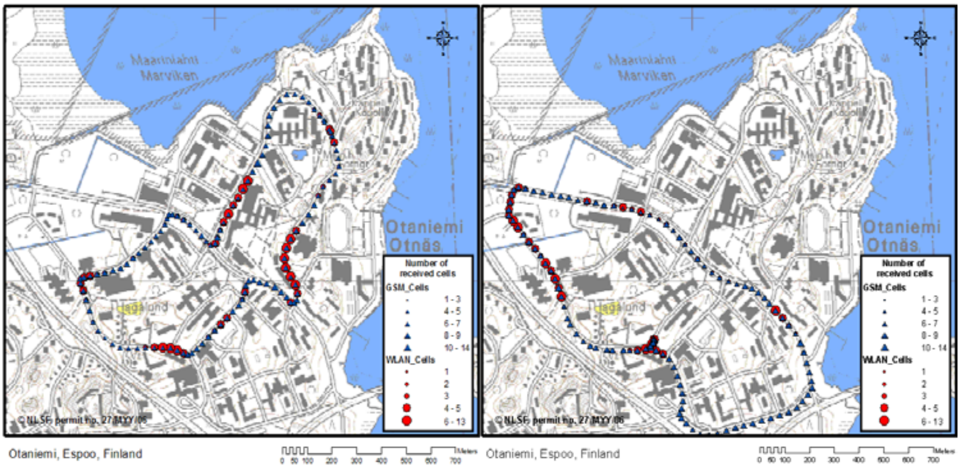


Figure 5. The distribution of received GSM cells and WLAN APs along the two test routes.

Subsequently, an evaluation of the positioning accuracy was performed by two separate test routes through the area (Figure 5). The first route was 4.0 km long and contained 114 samples. On average, 2.7 WLAN APs with standard deviation of 2.1 and 4.6 GSM cells with standard deviation of 1.8 were detected.

The length of the second route was 3.5 km and it contained 111 samples. In contrast to the number of detected GSM cells, which was equal to the first route, the number of observed WLAN APs was smaller. Only 2.2 WLAN APs were detected with standard deviation of 2.2. As can be observed from the figure, the second route contained more points with no WLAN coverage.

9.3.2 DCM Algorithm

The database concept originally proposed for GSM in [P1] was used in the positioning trial. The basic idea of the concept is to collect a database of reference signal fingerprints from the area where mobile terminals are to be located. The reference fingerprints are location-bound measurement samples, containing signal information from the detected cells. Instead of field measurements, the output from a network planning tool [R1] can be used to generate the database. Positioning is performed by matching the signal information of the request fingerprint, received from the mobile terminal, to the signal information of the reference fingerprints at LS. The location of the best matching reference fingerprint is selected to estimate the location of the mobile terminal. The concept of DCM is proposed for UMTS in [P2] and [P3].

The idea of exploiting data from multiple networks is presented in [P3]. Utilisation of data from both GSM and UMTS networks improves the positioning accuracy compared to the DCM resorting to only one network. In this chapter, the same concept is proposed for data obtained from both GSM and WLAN networks. In contrast to GSM and UMTS cells, WLAN APs usually reside indoors and their coverage areas are much smaller. The WLAN topology is also more varying compared to the GSM and UMTS networks: existing WLAN APs are shut down every now and then, or are moved to other locations and new networks are established in a short period of time. These characteristics have to be taken into account when extending the DCM algorithm to utilise WLAN APs. The following changes were made to the algorithm presented in [P3]:

1. Due to the smaller range of coverage, detected WLAN APs provide more accurate information about the mobile terminal location than GSM cells. Thus, it is worth using a scaling factor to emphasise the detected WLAN APs towards the GSM network data.

2. The topology of WLAN APs switched on may vary depending on the time of the day. In addition, new networks are established and existing ones are shut down quite often. Therefore, the penalty term both for missing request and reference WLAN APs should have a smaller effect to the correlation compared to the missing GSM cells.

9.4 Evaluation Results

The location estimates calculated with DCM were compared with GPS locations representing the true locations. The positioning accuracy is calculated as a Euclidean distance between these two location estimates. The positioning accuracy of DCM for the two test routes is shown in Figure 6 and in Table 2.

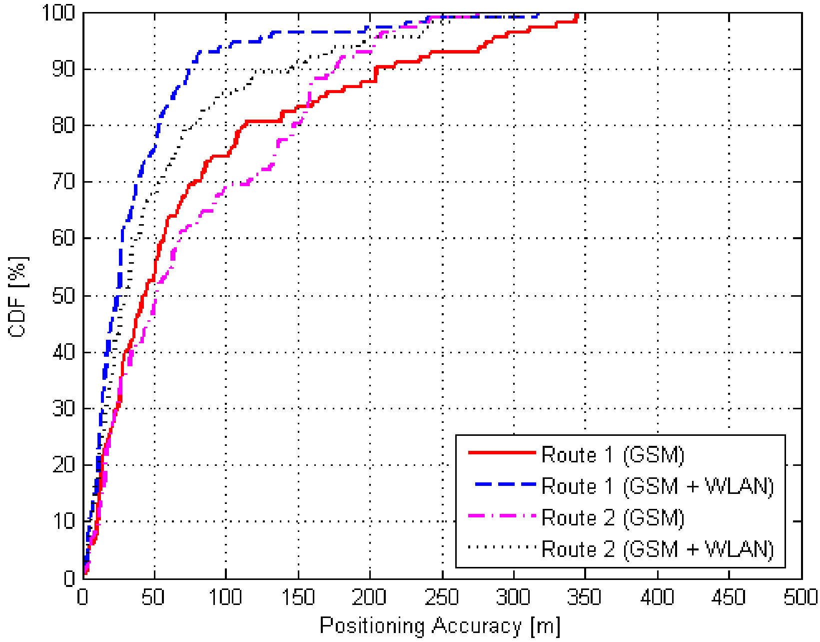


Figure 6. The positioning accuracy as a cumulative distribution.

Table 2. Summary of positioning accuracy.

Measurement	**67% value [m]**	**95% value [m]**	**Max error [m]**
Route 1 (GSM)	73	296	361
Route 1 (GSM+WLAN)	37	131	352
Route 2 (GSM)	98	207	303
Route 2 (GSM+WLAN)	52	234	307

When using reference data only from GSM network, 67% of the measurement samples were located within 70–90 meters from corresponding reference locations. Compared to the results presented in [P1] and [P3], the positioning accuracy was on the same level. The deployment of WLAN APs yielded better positioning accuracy along both routes. The 67% Circular Error Probability (CERP) varied from 37 meters to 52 meters with combined GSM and WLAN data. The more WLAN APs were detected, the greater the improvement was in the positioning accuracy.

9.5 Conclusions

In T1MGIS project, a lightweight, easily deployable measurement tool called WiNK was designed and implemented. It is capable of collecting signal information from GSM and WLAN networks. The data can be transferred from the measurement laptop to the server over wireless or wired TCP/IP connection. On the server side, the implemented MGIS system includes also a Location Server and MGIS Web Service that communicate with the measurement tool. The WiNK client can be used to assist LBS providers in verifying the prediction results of novel coverage estimation and mobile positioning methods.

In order to evaluate the software pertinence, a position trial was carried out in GSM and WLAN networks in a suburban area. Within the relevant coverage area, 188 different GSM cells and 357 WLAN APs were detected, and 2369 fingerprints were collected. Averaging the fingerprints decreased the number of fingerprints to 1795, which improved the performance of the positioning algorithm, the Database Correlation Method. DCM was extended by introducing the scaling factor and penalty term in the case of WLAN's involvement. Two test routes were analysed in evaluating the positioning accuracy. 67% CERP was reported at the range of 70–90m in GSM network, the same result as from the studies previously conducted by researchers. Yet with the additional involvement of WLAN, 67% CERP fell between 30 and 50 m. 95% CERP was within 300 m in the GSM network, and 240 m in the GSM&WLAN networks. Therefore, by utilising fingerprints collected from WLAN in addition to GSM network, the positioning accuracy was improved. Due to the fact of small coverage of WLAN APs, the more WLAN APs are present in the reference fingerprints, the more accurate the positioning can be. Most of the measurements were positioned in nearly half the distance from the true locations in the route with more WLAN APs than the one with less.

Besides in conjunction with GSM, WLAN based positioning can be applied independently in areas, which have complete WLAN coverage. With the wider availability of WLAN for most mobile devices in the future, and since most of the WLAN APs are concentrated indoors or in urban areas where GPS positioning encounters most problems, the authors believe the utilization of WLAN networks will be of help to perform sufficiently accurate positioning along with other cellular networks.

The current version of the MGIS software is a research prototype. However, it has already proven to be well applicable to mobile positioning development and multi-network signal fingerprint collection. Furthermore, the MGIS system will be heavily utilised in our research projects started at the beginning of year 2006.

Both WiNK and most of the other components in the MGIS system presented in this work have been implemented using Microsoft .NET technology [R10]. This choice of implementation enables Web service based measurement data sharing across different platforms but, in the same time, effectively restricts the MGIS system's own platform as Microsoft Windows.

Publications

[P1] H. Laitinen, J. Lähteenmäki, T. Nordström, "Database Correlation Method for GSM location." In: Proc. IEEE 51st VTC, 2001, Vol. 4. Pp. 2504–2508.

[P2] S. Ahonen, H. Laitinen, "Database Correlation Method for UMTS location." In: Proc. IEEE 57th VTC, 2003, Vol. 4. Pp. 2696–2700.

[P3] P. Kemppi, S. Nousiainen, "Database Correlation Method for Multi-System Positioning." In: Proc. IEEE 63rd VTC, 2006.

References

[R1] D. Zimmermann, J. Baumann, M. Layh, F. Landstorfer, R. Hoppe, G. Wölfle, "Database Correlation for Positioning of Mobile Terminals in Cellular Network using Wave Propagation Models." In: Proc. IEEE 60th VTC, 2004, Vol. 7. Pp. 4682–4686.

[R2] Nemo Outdoor. See http://www.nemotechnologies.com.

[R3] Tems Investigation. See http://www.ericsson.com/products/tems/.

[R4] Network Stumbler. See http://www.netstumbler.com/.

[R5] M. Wiacek. (2004, May). NetMonitor in Nokia Phones. [Online]. Available: www.mwiacek.com/gsm/netmon/faq_net0.htm.

[R6] L. Bursztein. (2006, Jan.). Fbus V2 Protocol. [Online]. Available: http://www.bursztein.net/articles/proto/proto.html.

[R7] Microsoft, (2006, Jan.). "NDIS – Network Driver Interface Specification." [Online]. Available: http://www.microsoft.com/whdc/device/network/ndis/default.mspx.

[R8] ESRI, ArcGIS. See http://www.esri.com/software/arcgis/index.html.

[R9] TeliaSonera. See http://www.teliasonera.com.

[R10] Microsoft .NET Homepage. See http://www.microsoft.com/net.

[R11] ArcMap in ArcGIS Desktop. See http://www.esri.com/software/arcgis/about/desktop.html#desktop_apps.

10. Open Hard Real-time Peer-to-peer Service Architecture (OHRA)

Tuomo Näyhä, Tommi Parkkila and Leila Rannanjärvi

10.1 Project Goals

The convergence of Internet and wireless communication domains has opened many new opportunities for the isolated information processing systems to become more effective and exploitable. This applies also to the scattered industrial automation systems, whether being a mobile work machine or a stationary production device. Both of these types of machine automation systems operate traditionally in isolation from the outside world for the strict safety and security requirements. A challenge lies in using the most modern technologies to create new kind of deeply networked, embedded real-time systems, which integrate machines and sensors safely and securely into the global information networks.

For the efficient networking of the demanding hard real-time automation systems – typically in the mobile machine domains – a new kind of service architectures are needed. A message oriented peer-to-peer (P2P) communication approach been found as a viable principle for connecting local information nodes to large global clusters in the Internet world. The P2P principle was initially chosen also here as a global networking mechanism. The other vital principles of the more locally acting architectural mechanisms were in the focus of this research.

The goal of the OHRA project was stated as “creating an open service architecture, where the power of the modern communication technology was to be harnessed to support the strict operational safety requirements of the hard real-time industrial automation systems”. The security issues, while challenging in P2P world, was left as a secondary topic in this research project.

10.2 The Course of the Research

The fundamental question of the OHRA project was how to apply the newest networking and data processing technologies to create new kind of deeply networked, embedded real-time systems. Deep networking means networking of distributed local (simple) embedded nodes into functional groups, driving computing deeper into the surrounding infrastructure. Looking from the global information networks, from Internet and wireless networks, the view should also be the same regardless of the size and function of the nodes. An utmost example of a deep networking is an internally and externally networked multicore system chip, where every node within a chip, whether virtual or real, expresses itself as a globally interacting entity that can join in any larger cluster to perform some higher level functions.

When analyzing the landscape of the automation systems, the first observation were the large amount of existing “standards”. which must be taken into account within any the new architectural system-level solutions. The approach here was an “integration architecture” (the OHRA architecture in Figure 1), where all existing (external) subsystems could be merged and new gradually applied (internally). The extent of the external subsystem could vary from a single sensor to a comprehensive real-time control system.

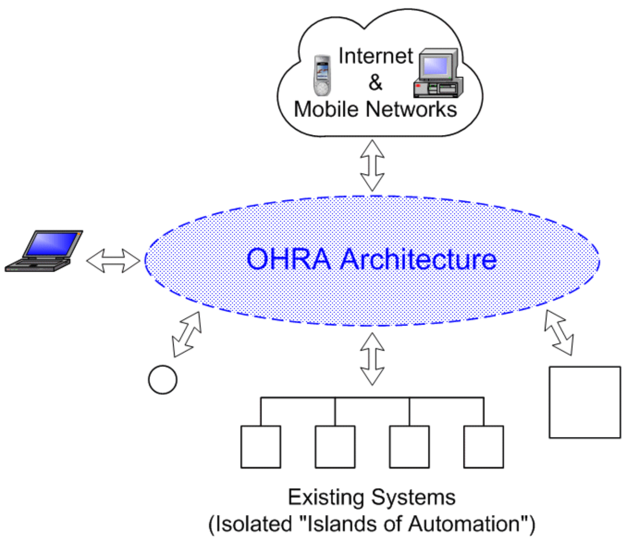


Figure 1. Architectural positioning.

From the beginning it was clear, that more powerful hardware platform than a conventional single chip controller was needed. In the year 2002 there were already available feasible configurable System-on-Chip (FPGA SoC) components, which were also economically suitable for the research platform and also for a small series of industrial application platforms. Accordingly, the FPGA hardware platform was chosen.

The architectural outlines could be sketched and implemented quite freely, because there were enough computational resources in an FPGA chip. Especially the flexible multicore architecture of the FPGA chip promoted the implementation of the deeply networked style of architectures.

As this research project was targeted to known specific systems (hard real-time systems) there were also thoughts to experiment the results in the real world. The building of the experimental hardware platform was vitally important for the evaluation of the functionality of the selected architectural solutions. There were also ambitions for further testing in some real-world application but it was withdrawn from this project phase due to the high cost of the needed external subsystems.

10.3 Description of the Results

The selected architectural principle was (Figure 2) building up local hubs (peer hubs), where a group of local real-time nodes were connected. For external subsystems, "the mirror nodes" were builded according the architectural outlines.

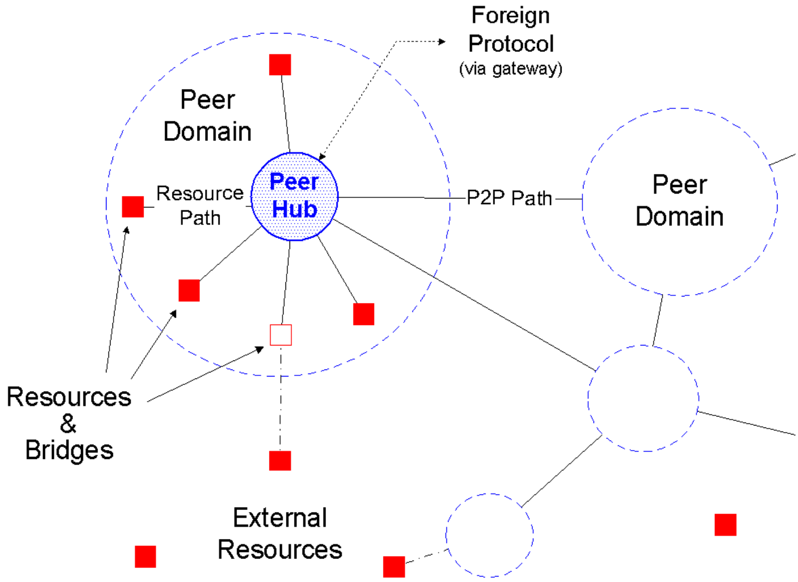


Figure 2. Architectural principle of the deeply networked system.

All communication between different node groups or to the global network was linked through peer hubs. In this way, a minimum amount of changes to the existing real-time nodes were required.

For the varying communication requirements in different parts of the system, the layered system architecture was selected. Initially the need for a tree layers (Figure 3) was identified: service layer, adaptation layer and binary layer.

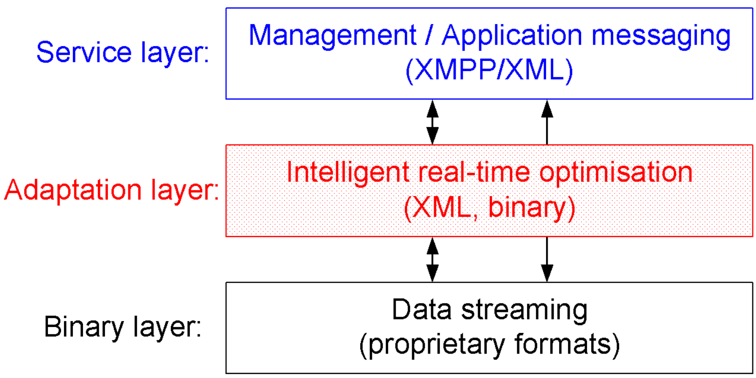


Figure 3. System layers.

The service layer was the most crucial in this architecture, because it enabled the global networking of the system. The lower level layers were ordinary real-time (or hard real-time) layers, which can only communicate within the real-time constraints of the underlying networks. So they were more or less local layers. The communication architecture that was chosen at the service layer was XMPP, which is also a widely used standard Internet protocol. This messaging architecture is by nature a hybrid P2P implementation, which had been found an effective and lightweight solution in earlier projects.

As each of the layers had different kind of communication requirements there can also be as many different middleware solutions. The Service layer message oriented middleware (MOM) architecture type comes directly from the XMPP system. Candidates for the middleware solutions for the other layers were also studied through literary research. The swarm intelligence was evaluated as the most promising mechanism to resolve the complexity of the distributed hard real-time system (adaptation layer).

The highest (service) layer was implemented fully as planned. Especially the embedded messaging system proved to be very useful and flexible. The implementation of the adaptation layer, crucial for the hard real-time activities, was left unfinished for the sake of the redirecting of the research theme in 2005. The implementation of the needed intelligence mechanisms are still somehow experimental and need further experimenting. These activities will continue in future research projects.

10.4 Conclusions

When integrating safety critical real-time systems efficiently into global information networks, new architectural layers are needed to preserve the integrity of the critical real time processes. In the course of this research project it became quite clear that the limited computational resources of the conventional real-time systems were the biggest obstacle for solving the problem only by adding new software layers to the system. The other hard problem was the diversity of existing real-time implementations and standards.

The selected approach in the project was to add new hardware resources to the system in the form of the system-level hubs and concentrating the new architectural elements only to these new hardware resources turned out to be successful. Mirroring the internal processes to this new architectural environment made it possible to manage real-time information safely without stressing existing real-time systems too heavily. The resources needed for the global networking were also isolated from the real-time processes and they were no longer a problem for the stability of the system.

A pathway to exploit this new architecture can go smoothly first by integrating existing real-time systems into it and then gradually building new solutions according to the new architecture outlines. By preserving existing investments of the often very costly control systems, the acceptance to the new system-level solution is much better than if discarding the old systems were a precondition.

The most significant achievement of the project was the definition and implementation of the embedded messaging system that complies with Internet communication standards and can penetrate deeply into this new system architecture. The possibility to manage and control deeply into system embedded real-time functions remotely through global information networks had enabled the creation of a wholly new class of real-time applications.

The results of the OHRA research project have been successfully used in several application oriented projects. Also new research activities have emerged directly after the project was closed in 2004. Locally the OHRA project has successfully acted as a starting point for a new field of research.

Publications

[P1] L. Rannanjärvi, T. Näyhä, "The Service Architecture for Mobile Machines and Flexible Production." In: Proceedings of the ICMA02, the fourth International Conference on Machine Automation, Human-friendly Reliable Mechatronic, Tampere, Finland, Tampere University of Technology. Sept. 2002. Pp. 435–445. ISBN 952-15-0861-2.

[P2] T. Parkkila, "Remote management of an embedded system on system-on-chip." University of Oulu, Finland, Diploma thesis, 2005.

[P3] T. Parkkila, "Application and platform management of an embedded system." Smart Systems 2005, Conference Proceedings, Seinäjoki, 3–4 May 2005.

11. MEMS Radio

Markku Åberg, Ari Alastalo, Jouko Aurinsalo, James Dekker, Jaakko Lenkkeri, Hannu Salminen, Heikki Seppä, Stefan Tallqvist and Miikka Ylimaula

11.1 Introduction

Development of process technology has made it possible to manufacture also mechanical devices in addition to electronic ones with the size of microelectronic components – in the order of micrometers, thus the name “micromechanics” or “microelectromechanical systems” (MEMS). This development has made it possible to make mechanical components

– with new functions
– for parameter ranges impossible before
– integrable with microelectronic circuits
– mass produced with very low price per device.

This development has generated the idea of using micromechanical devices in high frequency systems, for instance in radio transmitting and/or receiving systems.

11.2 Objectives

The original objectives of the project were as follows: “The goal of the project is to develop a short-range radio transceiver, based on micromechanical components. The architecture of the radio will optimized to utilize optimally the properties of micromechanical components. The number of these components will me maximized. The micromechanical components needed will be developed in the course of the project.”

In the course of the project the different MEMS building blocks, mainly as discrete components, were seen as the critical items. The status of the technology does not yet allow a monolithic integration of these components. VTT has made substantial progress in developing monolithic integration of MEMS and CMOS, but the development has been parallel to this project, and the results therefore were not even in principle available for this project before the end of activities, and on the other hand the demands of this project for the parameters have been so strict that only discrete components have had some hope in fulfilling them.

The project was divided into the following tasks, defining the actual activities:

Radio Architectures

The goal was to find the architectures that are the best for MEMS radios. Also totally new architectures based on micromechanics were to be studied.

Micromechanical Components

The goal was to develop micromechanical radio frequency devices. The devices could already exist in principle, e.g. switches, resonators, varactors, but also totally new types of devices would be studied. The developed (designed) devices were to be processed and evaluated. No generic process development was planned to be done in this project. Instead, development work done in other VTT projects was to be utilized

Packaging

The goal was to develop an optimal package for a MEMS radio. The first target was to find a functionally optimal solution, but also cost and production aspects had some weight.

Antennas

The goal was to develop miniature antennas, hopefully of such kind that they could be integrated with the packages.

Integration

The goal was to maximize the integration level. The first task was to develop modules with MEMS devices and integrated electronics. The final task was system integration.

Demonstrators

The goal was an operative demonstrator of a whole radio system with a selected architecture.

BAW Devices and Circuits

The bulk acoustic wave (BAW) devices and circuits development was an independent project with the name "New Filter Technologies" the years 2002–2004. In 2005 this project was merged with the MEMS-Radio project. This (sub)project has a separate chapter, chapter 13.

11.2.1 MEMS Radio Architectures

MEMS based components have been suggested to implement many of the needed functionalities in modern communication devices. A typical superheterodyne radio architecture is depicted in Figure 1 ([R1]). The blue shading indicates components that can possibly be realized with MEMS technology. In addition to the switches, phase shifters, local oscillators, filters and impedance tuners that are generally recognized as potential applications of MEMS ([R2]–[R9]), for the power sensor, used in the radio of Figure 1 to control the power amplifier (PA), one can consider using MEMS, as first discussed in [R10] and developed further, for example, in [R11] and [R12]. Instead of developing MEMS replacements of conventional components, full utilization of the properties of MEMS may well require a redesign of the overall architecture and may prove well suited also for low-power low-performance radios of wireless sensors. Namely, MEMS may allow some of the different operations of Figure 1 to be performed in a single electromechanical device, such as the receiver front-end filtering, mixing and even the low-noise amplification using parametric pumping [P11] and [R13]. In [R6] and [R14], an architecture is considered, where the receiver (RX) band-select filter is replaced with a switchable bank of narrow-band filters with different center frequencies to cover the receiver band. With such an approach, the linearity requirements of the low noise amplifier and the mixer can be relaxed due to the introduced selectivity against in-channel interferences. A multitude of parallel channel-select filters is also considered in [R15] with bulk acoustic wave filters for a radio of wireless sensors. Instead of many parallel filters, one can also consider using a single multimode MEMS resonator, such as the plate resonator of [R2] with Lame and square-extensional modes, if the size of the radio is to be minimized.

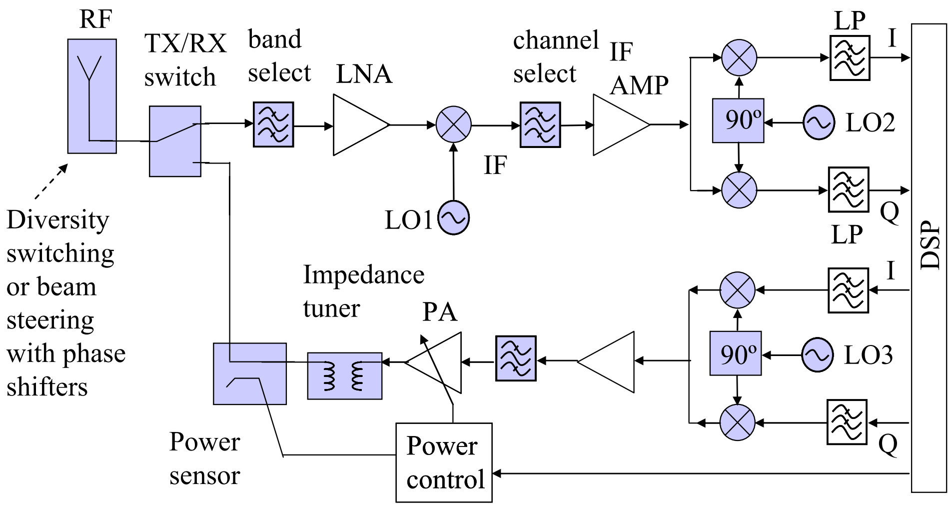


Figure 1. Superheterodyne radio architecture. Components for which MEMS implementation can be considered are indicated in blue.

Figure 2 schematically shows another possible architecture for a low-power radio transponder terminal that communicates with a high-power base station. Here the transmitter carrier is generated from the receiver signal using a time delay during which the transmission/receiving switch changes its state. The carrier generation can comprise of for example a phase-locked loop or, if the base station is sending an unmodulated carrier for the terminal transmitter, pass band filtering can be enough. With a narrowband delay line, the filters in Figure 2 can possibly be omitted. The spring-mass-chain delay line that is developed in the project has such a narrow-band response and could be used in a low-power transponder as schematically shown in Figure 3 ([P5] and [P10]). Here the RF pulse sent by the reader is either echoed back (bit 1) or shunted to ground (bit 0). Surface acoustic wave RFID tags are another well-known example of a commercial narrow-band transponder radio, where the input reception pulse is coupled to a delay line with multiple reflectors that cause many transmission pulses, with identifiable inter-pulse delays, to be sent back.

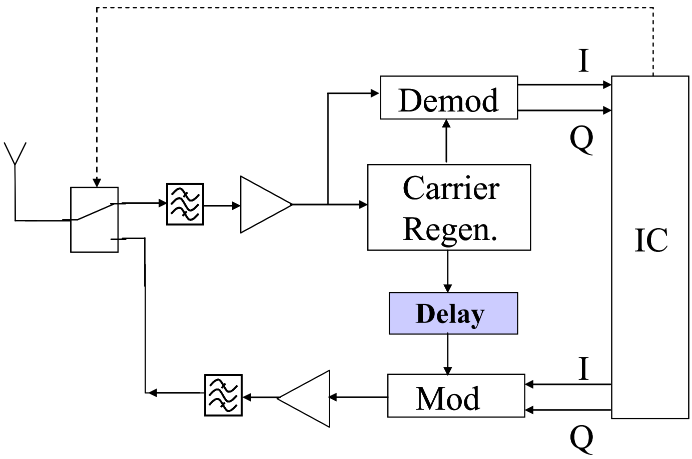


Figure 2. Schematic of a radio-transponder architecture where the delay line is used to separate reception and transmission in time.

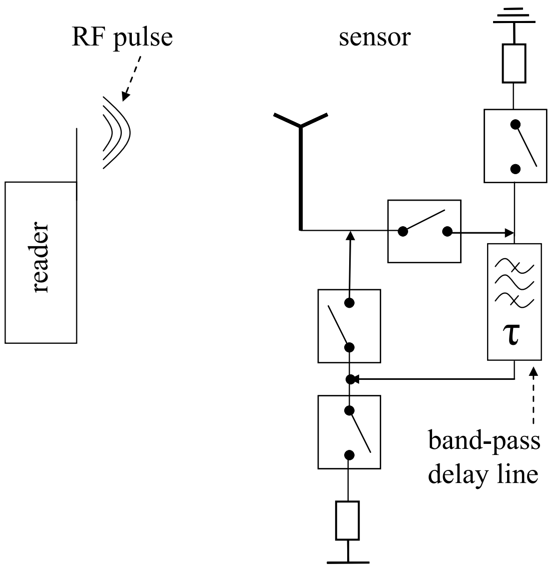


Figure 3. Schematic of a simple delay-line transponder for low-power low-data rate communication in sensor applications [P5].

11.2.2 Process Developments

The wafer processing done during the course of MEMS radio followed two routes: 1) standard processing, and 2) new process development. The standard process for MEMS resonator fabrication, well-known from earlier VTT reports [P11], was used to fabricate devices as needed for testing designs. In addition, selected designs were chosen and fabricated using the new processes developed during the project. The goal of the new processes was to provide both wafer level packaging and narrow resonator gaps to improve resonator performance. Furthermore, the new process was intended to circumvent certain problems which occasionally arose when the standard process was being implemented-namely, metal residuals which may hinder gap etching if it is performed after metallization. Although several process runs were made with both types of process, this document will discuss only new process development.

The new process, labelled Membrane for convenience, is illustrated in Figure 4. It is most notable in that it utilizes direct wafer bonding after the gaps are etched, followed by stripping of the cover wafer to reveal the resonators beneath a 1 μm oxide, later used as the basis of a protective membrane. Metalization is performed last and the devices may be sawed (using protective tape). Not shown in the figure 4, is the possibility of incorporating a gap narrowing process before gap etching. This will be described later.

Figure 5 shows a first generation device fabricated using the illustrated sequence with a membrane cover over a low frequency (750 kHz) mixer. The operation of the device (uncovered) has been presented elsewhere [P10]. With first generation devices it was possible to measure resonances. However, stress concentrations around the membrane edges lead to frequent cracking of the membranes.

Figure 6 shows a second generation device. The refined process makes several notable changes from the first generation. Firstly, the cover wafer was stripped, rather than the SOI handle layer as shown in Figure 4-c. Secondly, the cover defining lithography shown in step 4-d was skipped. Third, an oxide mask was used rather than PR mask in step 4-a, to facilitate the gap reduction process described later. Fourth, a second window lithography and etch step was added after step 4-i to facilitate window etching through the membrane (this is optional).

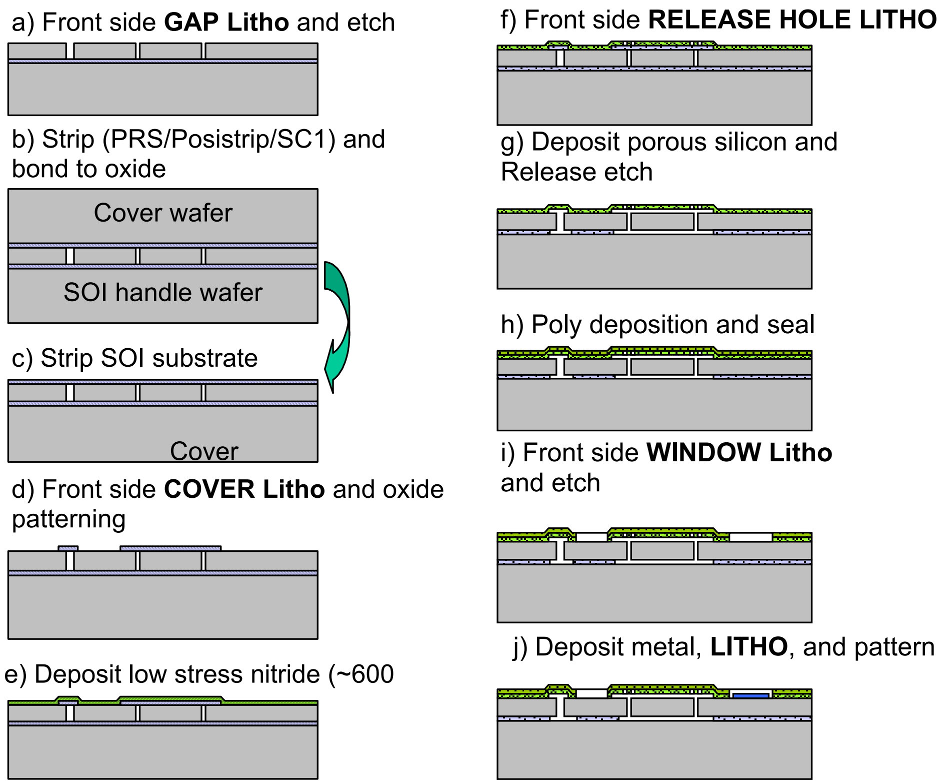


Figure 4. The process flow for the membrane process is shown here. 1) resonator lithography and etching followed by resist removal. 2) wafer cleaning and bonding, 3) substrate stripping, 4) cover lithography and oxide etching 5) nitride deposition, 6) lithography and etching of release holes through the nitride, 7) porous polysilicon deposition, release etch and SC drying, 8) poly deposition to seal the cavities, 9) window lithography and patterning (to reach SOI device layer) and 10) final metallisation and patterning of metal pads.

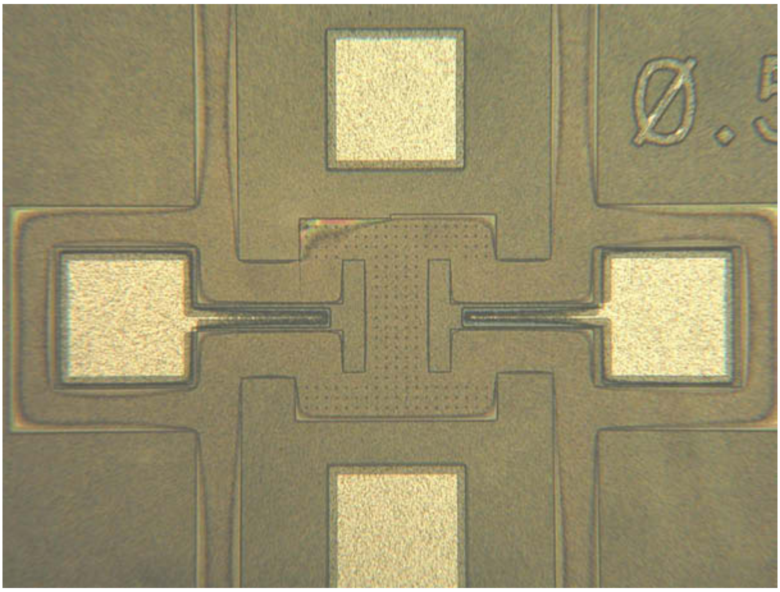


Figure 5. A mixer beneath the first generation membrane cover. The well defined cover unfortunately suffered from cracking at its edges.

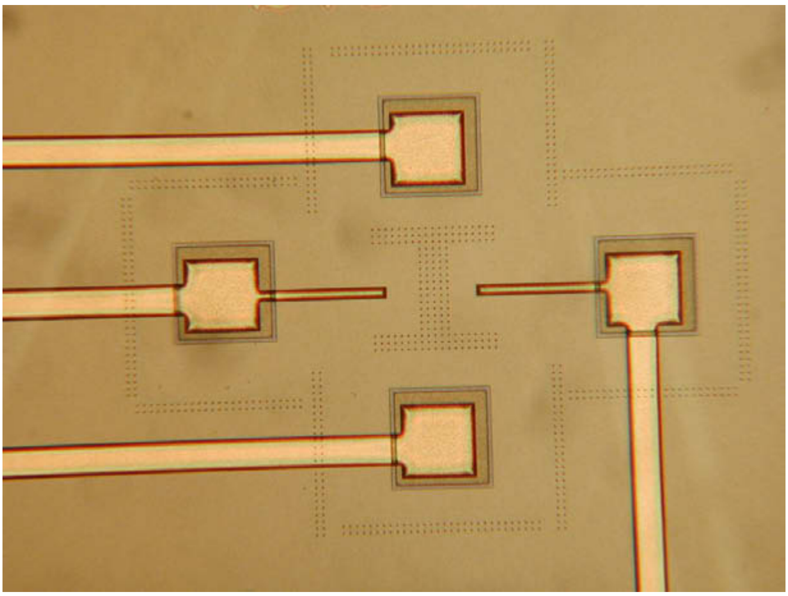


Figure 6. This shows another mixer structure beneath the second generation membrane cover, which eliminated cracking and allowed more extensive metallization. This was achieved by simplifying the process and eliminating a mask level.

The two most important changes are the elimination of the cover litho and the stripping of the cover wafer, rather than the Silicon-on-insulator (SOI) substrate. Removing the cover litho eliminates the high stress at the membrane edge and therefore reduces the tendency to cracking. This virtually eliminates cracked membranes from the process. Note that the decision to eliminate the membrane defining litho could only be made after the first process run, where it was observed that the etch rate of the bonded interface (below the membrane) was comparable to the buried oxide etch rate. This relies on high quality bond strength only achievable with fusion bonding. The cleaning procedures required to achieve such results involved resist stripping in the PRS800 Oxygen plasma stripping tool plus the posistrip bath, (or in some cases mask oxide removal in HF), followed by an SC1 clean. In some cases an Argon plasma was added as a pre-bond activation, though this had no clear benefit. Bond annealing was done at 1050 °C. Figure 7 shows a scanning acoustic microscope image indicating the good bond quality and Figure 8 shows an infrared (IR) image in which the bond fronts for the two oxides are seen to coincide. Figure 9 shows a set of longer delay lines. Even in such large delays lines the membranes are intact although there are now occasional “bubbles” formed by weak spots in the bonded interface. The origin of these weak areas (which will have a higher oxide etch rate), may lie in the highly doped nature of the wafers, as such wafers are known to sometimes exhibit small voids following bonding. Some of these spots are actually visible in the scanning acoustic microscopefimage of Figure 7, in which they are distributed in rings near the wafer edge. Note that this is a random occurrence and about half of the wafers were almost perfectly bonded without such voids. Despite the fact that the IR inspection showed the devices were released and the membranes intact, these long delay lines were not functional.

Three different methods of removing the cover wafer were tried. Initially, this was done with an isotropic SF6 plasma in the ICP etcher. The etch time to remove the 380 μm thick SOI substrate was less than 2 hours. This has the benefit of being very gentle and the etch has good selectivity to stop on the oxide. The next preferred method was grinding. It was found that if grinding was used to remove all but the last 30 μm, then the final material could be removed with isotropic SF6 in just 20 minutes. Typically, however, 50–80 μm was left after grinding. Finally, an obvious alternative was to use TMAH. This takes roughly one day but can be done in batches. However, it was found that the oxide membrane tended to break in the TMAH bath and several wafers were lost this way. Thus, the preferred method for stripping was grinding to ~50–80 μm followed by plasma etching.

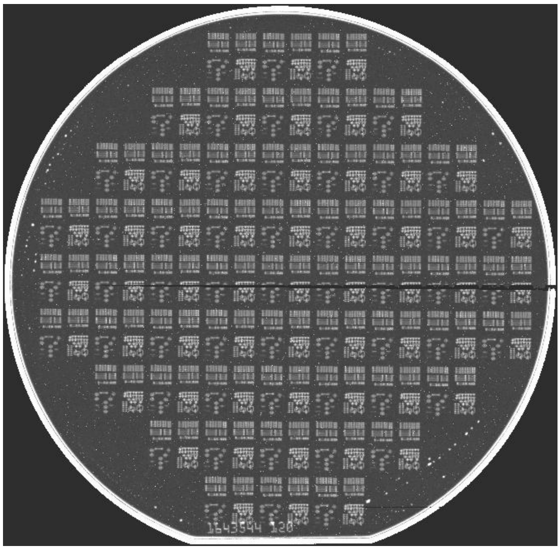


Figure 7. A Scanning Acoustic Microscope image of a bonded pair after annealing. The circular pattern of spots near the wafer edge are due to the heavy boron doping. In this case, bonding used both SC1 plus plasma activation. Megasonic cleaning immediately before bonding was done at reduced power.

The second process development made during the project was a method for reducing the capacitive gap width of the resonators. The process uses the standard lithography tools of the clean room to make the initial pattern which is etched through the oxide mask and then roughly 1.5 μm deep. After cleaning, the shallow trench is then oxidized so that the sidewalls close together, reducing the effective width. It is then necessary to perform a short plasma etch of the oxide to open the trench bottom and then the main silicon trench etch. During test procedures to evaluate the process, it was possible to make trenches down to 250 nm width. During the first process run, the goal was to get a 300 nm trench through 5 um SOI with a final aspect ratio of 16:1.

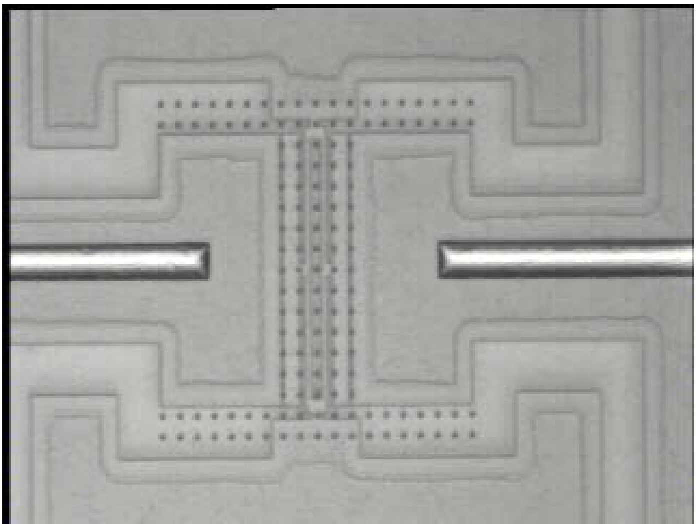


Figure 8. IR image of the released resonator. The etch front for the buried oxide and the upper membrane oxide both coincide, indicating the high-quality bond strength.

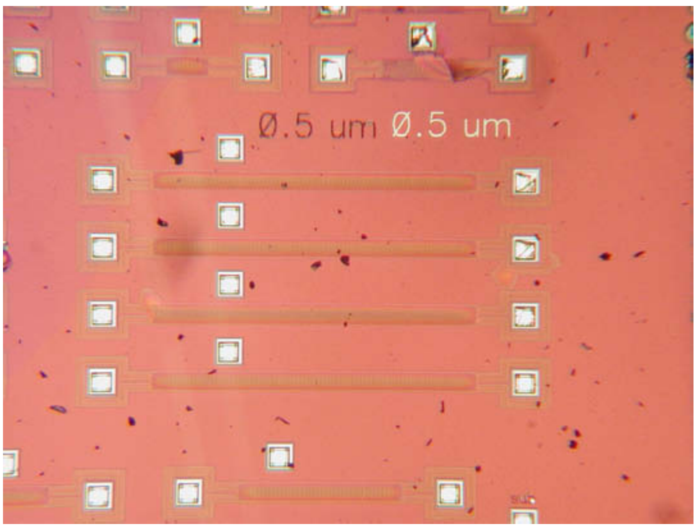


Figure 9. Long delay lines after the final process. The membranes are intact, but the small "bubbles" beneath the membrane show where voids were present following bonding of some of the wafers. These are presumably the same voids due to the p-plus doping seen in the SAM image of Figure 7.

The first attempt to combine narrow gap resonators with membrane covers was moderately successful. Unfortunately, an error in mask fabrication meant that the narrowest structures could not be fabricated, but slightly wider structures were still made. The gap narrowing process successfully reduced 0.9 µm trenches to below 0.6 µm width, as demonstrated by electrical measurements. On the same process, 0.5 µm structures, which were targeted to be 250 µm wide, were also etched but, due to the aforementioned mask error, the gaps contained "shorts". Nevertheless, some structures were etched and optical inspection shows the gap narrowing process and etching was successful, though no operating devices were obtained.

One difficulty of combining the gap narrowing process with the membrane process is the need to remove the residual oxide etch mask before wafer bonding. This requires a short etch in HF and it must not break through the buried oxide if is to be used for the membrane as shown in Figure 4 (note that figure 4 illustrates the standard process with a PR mask – for the gap narrowing process, an oxide mask would be used in step 4-a). Fortunately with proper process design the residual oxide can be 200 to 400 nm after the trench etch and this is removed in just 5 sec in HF without losing the buried oxide. If the device layer is not "flipped" by substrate stripping, stripping the cover wafer instead, then it is not critical that the buried oxide remain intact- this change was incorporated into the second generation process.

Figure 10 shows the resonance curves measured on the parameter analyser. The 0.9 µm lines without gap reduction could be seen to resonate with a bias voltage of 15V had a pull-in bias of roughly 40V, whereas the narrowed devices showed resonance at 7V and had a pull-in of less than 10V. This indicates a significant reduction in line width, possible as low as 500 nm, but more realistically in the range of 600 nm. Based on monitor data, this process can be scaled to still narrower line widths below 400 nm. It should again be emphasized that this is accomplished with standard lithography in an i-line stepper.

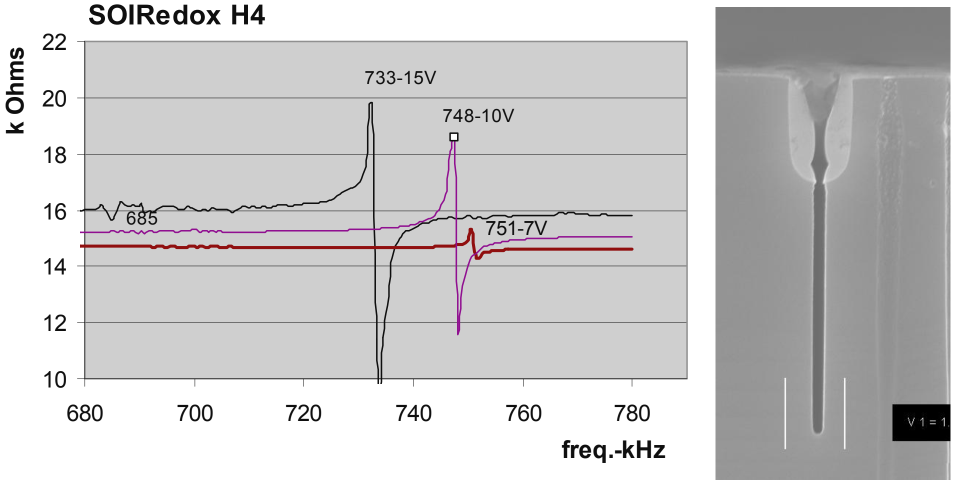


Figure 10. Resonance curves made on the parameter analyser for mixers fabricated with reduced gaps and wafer level encapsulation. For devices with reduced gaps, the resonance began at 6V and pull-in occurred below 20V. For devices made without the gap reduction process, pull-in did not occur until over 40V. A narrowed gap is shown on the right. The scale marker is 1 µm. Developments have been made to reduce the size of the oxide sleave at the top of the trench. In an actual resonator, the oxide is removed during the etching step, and the process does not appear to have a detrimental effect on the resonator.

To summarize, the processes developed within MEMS Radio showed the capability to provide both wafer level packaging and narrowed electrode gaps for SOI resonators. The process to accomplish thus is relatively simple and requires no extensive deposition or arduous etchback/polishing operations. The most demanding step, the wafer bonding and thinning, is done early in the process which minimizes the chance of failure at that stage. Electrical testing confirmed the operation of the devices, although the yield was low. Nevertheless, process modifications have been envisioned which could alleviate this difficulty.

11.2.3 MEMS Delay Lines

Usage of electrostatically-actuated longitudinal waves in a silicon rod for signal transmission, as schematically shown in Figure 11, has been analyzed in detail [P4]. For practical realisation of this kind of MEMS waveguide, impedance matching is a challenge. This is because the weakness of the capacitive coupling makes the characteristic electrical impedance of the line extremely high.

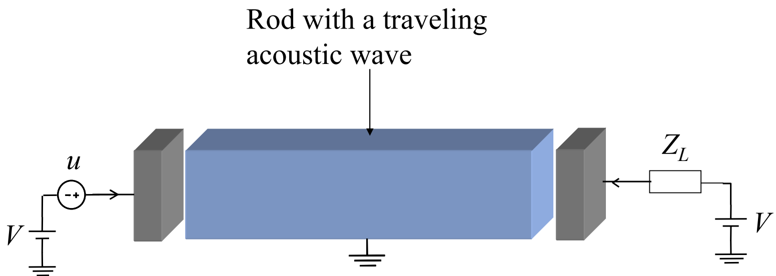


Figure 11. Schematic representation of a setup where an electric signal is transmitted through a micromechanical rod as an acoustic plane longitudinal wave.

For HF frequencies, a new kind of delay line that is based on a chain of MEMS resonators has been developed ([P5] and [P10]) in the project. A scanning electron microscope (SEM) micrograph of one of the fabricated devices is shown in Figure 12. The resonator-chain delay line is characterized by extremely slow signal group velocity (~ 10–100 m/s), narrow-band response, and much lower characteristic impedance than found for the solid-rod waveguide enabling efficient signal coupling. The narrow-band characteristics can be utilized in applications that would otherwise require a separate bandpass filter as in wireless transponder radios of sensors for which a possible structure was suggested in Figure 3.

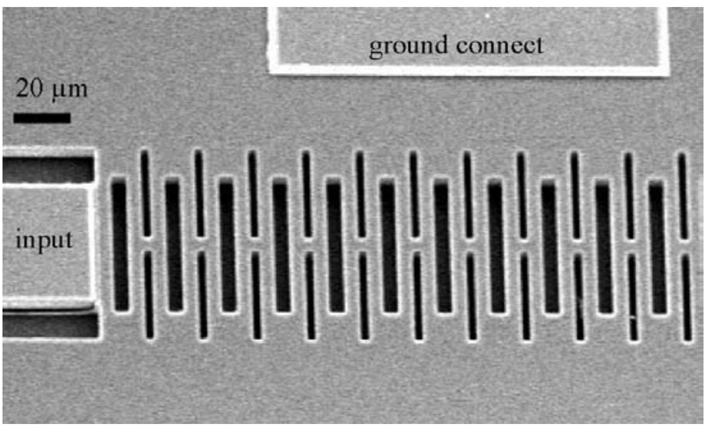


Figure 12. SEM micrograph of a fabricated spring-mass-chain delay line.

MEMS Filters

In this project, intermodulation properties of capacitively-coupled bandpass MEMS filters were analytically solved in closed form and the trade-off between linearity and insertion loss was quantified [P1–P3, P12]. The theory was first formulated for weakly-coupled single-stage filters and then generalized to strong coupling (low motional resistance and low insertion loss) and to higher order filters. Usage of a single-stage MEMS resonator as a filter is shown in Figure 13.

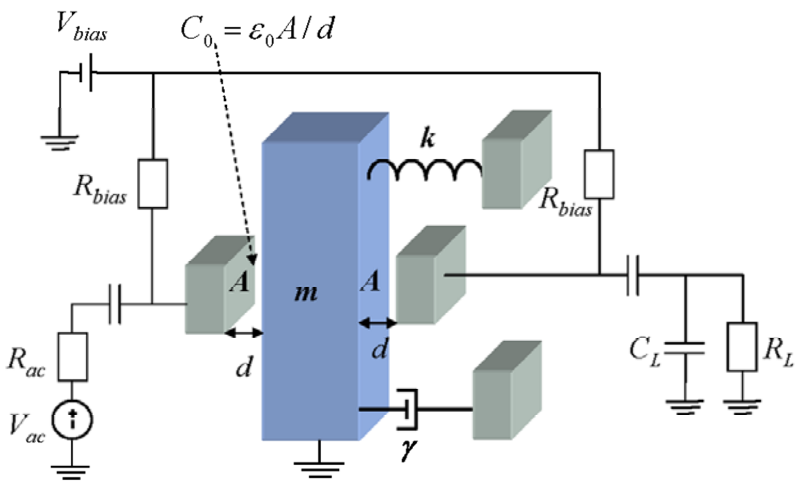


Figure 13. Usage of a MEMS resonator as a filter.

The theoretical results were verified in circuit simulations as well as in measurements. Also mechanical nonlinearities were included although, typically for good coupling, the capacitive nonlinearity is the dominant source of intermodulation. Figure 14 illustrates the intermodulation distortion as a result of third-order nonlinearity in the system.

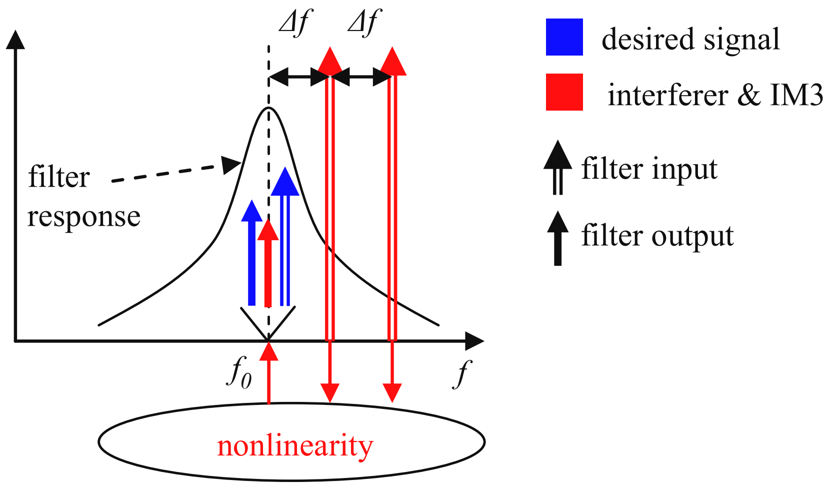


Figure 14. Nonlinearity as a source of intermodulation.

The obtained formulas are more generally applicable than the previously published results ([R16] and [R17]). After the analysis, the theory is utilized to formulate a design procedure for MEMS filters that, for the first time, takes systems specifications for tolerable intermodulation distortion and insertion-loss into account ([P3], [P12]). The conventional resistive 50 ohm source and load termination was shown to typically result in a high insertion loss if good linearity is required with low-voltage operation. Consequently, it was found desirable to utilize the high resonator quality factor for voltage gain that is enabled by capacitive load termination at the filter output. This is possible in integrated receiver architectures, where 50 ohm transmission lines are not needed between the antenna and the filter and between the filter and the low noise amplifier as shown in Figure 15. With such an approach, MEMS filters could be used to construct a receiver front-end having a bank of narrow-band (ultimately channel-select) filters with different passbands to cover all the receiver channels as suggested in [R6].

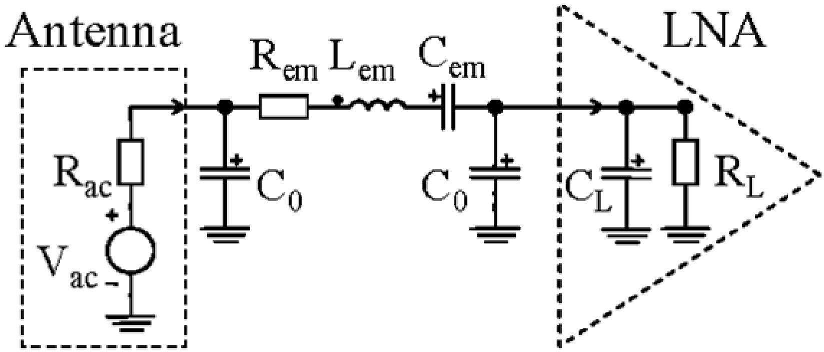


Figure 15. Integrated MEMS radio front end.

11.2.4 Small Antennas

An antenna in general efficiently converts electromagnetic energy to spatial radiation. However, if the dimensions of the antenna are small in comparison to the radiating wavelength, the efficiency of the antenna will decrease considerably and the radiation will be essentially non directional. The radiation characteristics depend also much on the dimensions and properties of objects in the immediate surroundings of a small antenna. Even a small antenna has to be fairly well matched to the feeding source. The feed impedance control of a small antenna is a major design problem. The success in this matching task has a strong effect on the radiation efficiency.

Many types of small antennas, generally embedded in some ceramic material, are today commercially available. The properties of such antennas, e.g. radiation characteristics and feed impedance, are fairly well known from figures provided by the manufacturer. Therefore, these commercially available small antennas were not investigated during the project.

Promising new type small antenna designs were investigated both theoretically by using simulation software and also by practical measurements in an anechoic chamber. Because neither the exact operating frequency nor the design of the final device were known, most measurements were made on a test board, equipped with a SMB connector and a stripe-line feed. The measurements were made on frequencies around 0.8 and 2.4 GHz.

The dimensions of an antenna might be made smaller by shaping the equivalent length of a straight radiator to a coil or meander line shape. The physical dimensions might also be made smaller by use of some ceramic material with a high dielectric constant. This ceramic material can be located close to the conducting antenna structure, or the antenna can be entirely embedded in the ceramic material. The high dielectric constant ceramic material may also be made to resonate and radiate by itself. A slot in the conducting housing of a small electronic device may also be made to radiate by careful design.

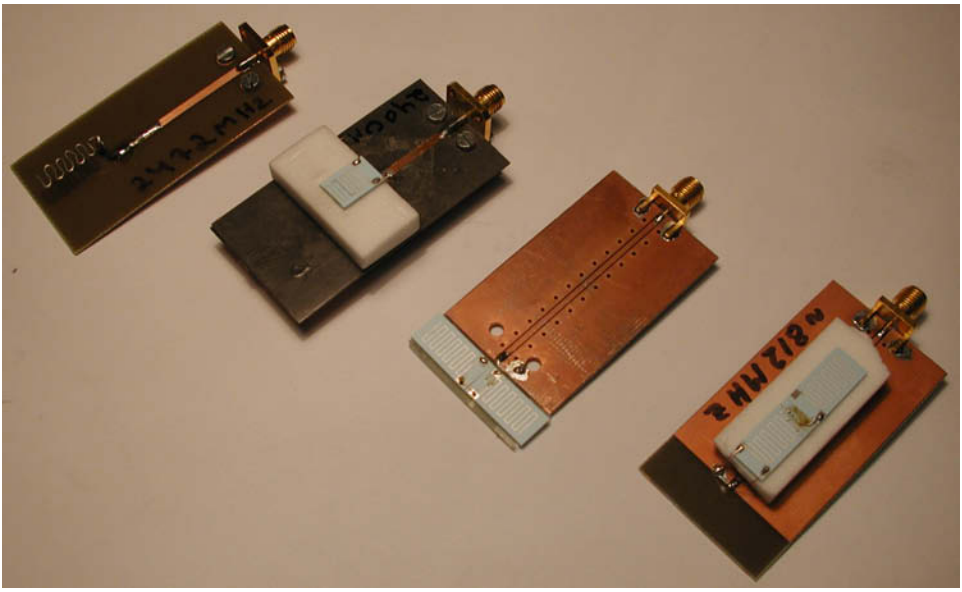


Figure 16. Examples of small meander antennas investigated both theoretically and experimentally.

Investigated and measured antenna designs were of the following types:

- a: small meander-line antennas in various configurations and with different antenna feed structures
- b: small ceramic patch antennas with various feed structures
- c: small antennas consisting of a flat conducting coil with various designs.

Examples of small meander-line antennas investigated both theoretically and experimentally are shown in Figure 16. It was generally possible to find a fairly good impedance matching for this antenna type, both with and without a ground plane beneath the radiator. A simulation model of a meander-line antenna is shown in Figure 17 a). If the test board had a ground plane beneath the meander-line conductor, a zero point appears in the antenna radiation pattern in the direction beneath the ground plane, see Figure 17 b). Measured gain values of such meander-line antennas were around 0 dBi.

Ceramic path radiators with various feed structures, for example a slot feed from a strip-line, were investigated. This feed type was more difficult to simulate compared to the case of the meander-line antenna. The resulting bandwidth was often narrower in practice. An example of this antenna type is seen in Figure 18. Measured gain values and efficiencies for these small antennas were generally lower than for meander-line designs.

Small antennas consisting of a flat conductor shaped as a spiral coil were investigated, see Figure 19. These conducting coils might possibly be embedded in some ceramic material. Considering the easiness of the design and the measured characteristics these coil antennas were comparable to the small meander-line antenna designs.

The mentioned small antenna designs were investigated on a general level only. If a specific electronic circuit is embedded into a small LTCC (Low Temperature Co-fired Ceramics) ceramic housing, these small antenna designs need to be developed further. In many cases the best solution may be to use some commercially available small ceramic radiator. Some method to transfer electromagnetic energy from the electronic circuit to the antenna radiator has to be found in each situation.

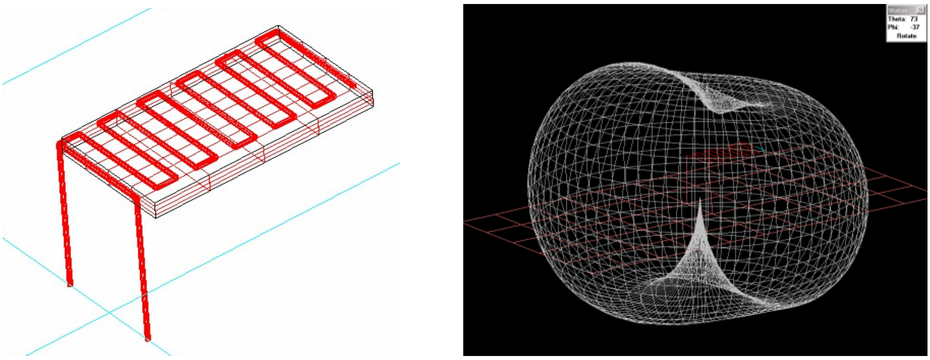


Figure 17 a) Simulation model of a small meander antenna for theoretical calculations; b) Calculated 3D radiation pattern of a small meander antenna on a test board.

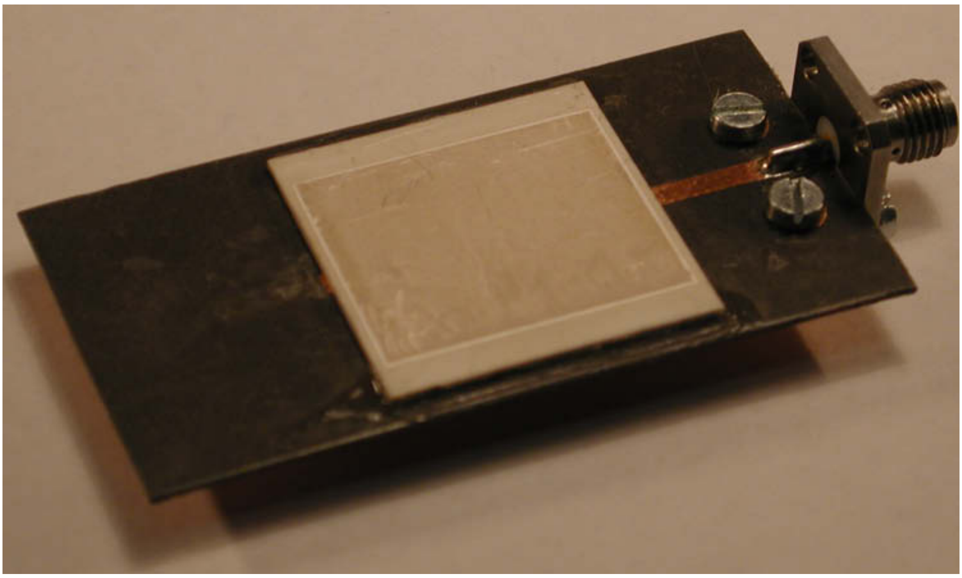


Figure 18. A ceramic patch radiator fed by a slot from a strip line.

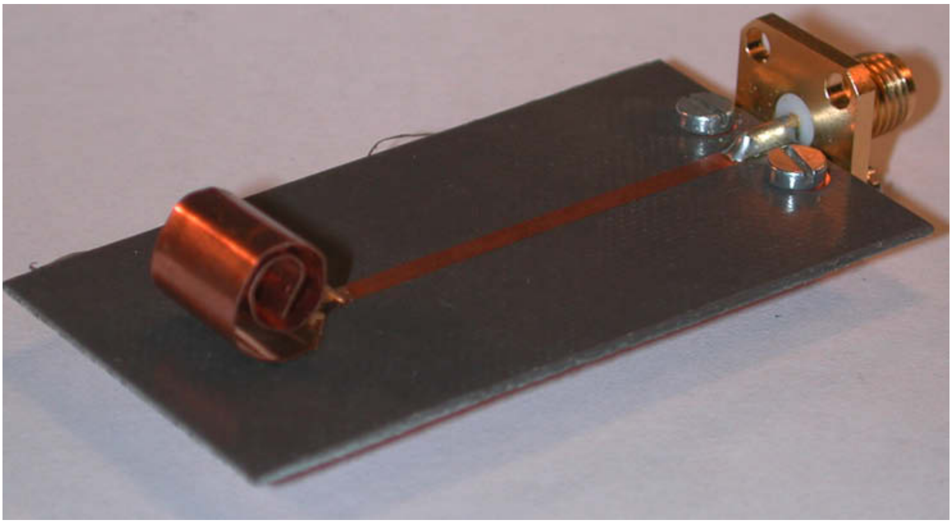


Figure 19. A small spiral antenna measured on a test board.

11.2.5 Bulk Acoustic Wave Device Integration

The integration development was concentrated on oscillator circuit blocks. MEMS varactor and varactor-based circuits research was moved by the end of 2004 to other projects, e.g. ESA/ESTEC project "MEM Capacitances in Delay Lines, Transformers and RF Signal detectors". Thereafter this project concentrated on Bulk acoustic wave (BAW), also called as Film bulk acoustic resonator (FBAR), based oscillator and clock circuits.

The FBAR oscillator presented here was designed for a data-converter application in which low noise 300MHz master clock with differential output was required ([R18]). To meet this specification we used 0.9GHz FBAR oscillator, divider and single-ended to differential converter. System block diagram is shown in Figure 23.

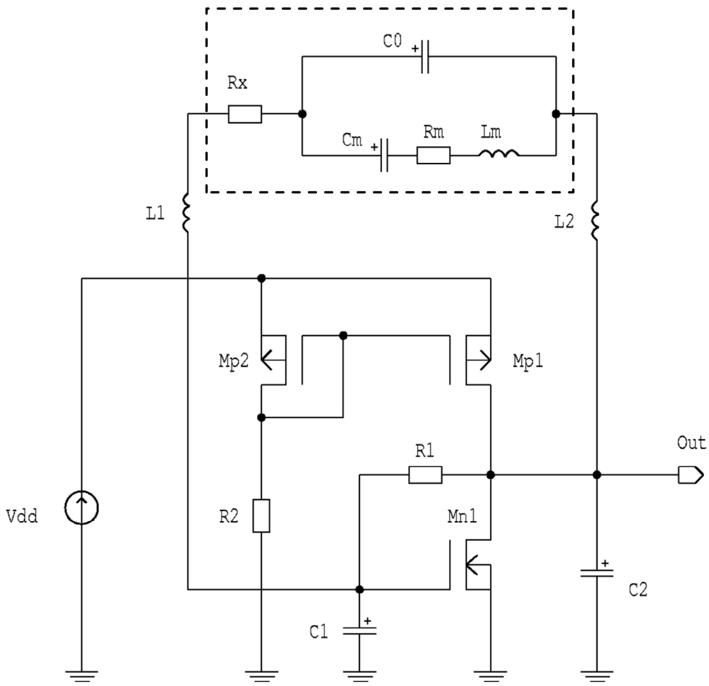


Figure 20. Oscillator circuit, the box with the broken line shows the FBAR model.

The oscillator circuit consists of common source amplifier and positive feedback network formed with FBAR and capacitances C1 and C2 (Figure 20). This circuit configuration is the well-known Pierce oscillator or three-point oscillator.

Inductances L1 and L2 are bond-wires between the FBAR and the amplifier chip. Transistor Mn1 provides transconductance needed to generate and sustain oscillations. The rest of the circuit, PMOS current mirror (Mp1–Mp2) and resistors (R1–R2), is a bias network used to set transistor Mn1 to proper operation point.

The FBAR model consists of a motional impedance (Rm-Lm-Cm) and a plate-capacitance C0. Motional resistance Rm models acoustic attenuation of the FBAR. To generate oscillations amplifier gain must be greater than the acoustic attenuation of the resonator, in other words the negative resistance seen by the motional impedance must be higher in magnitude than Rm.

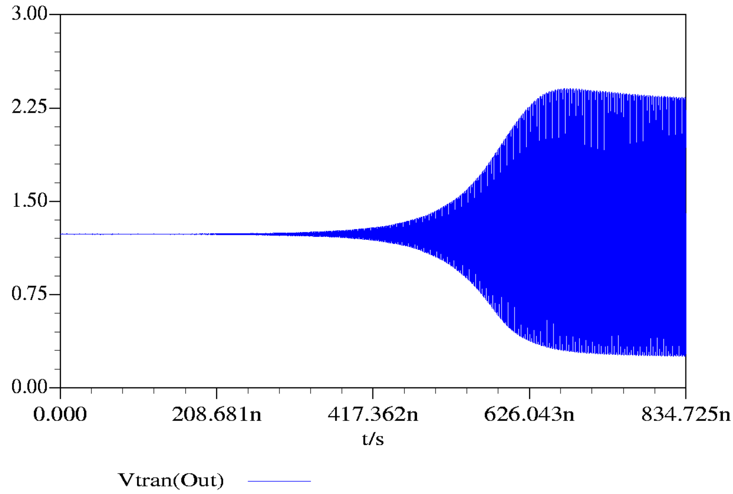


Figure 21. Simulated start-up of the oscillator.

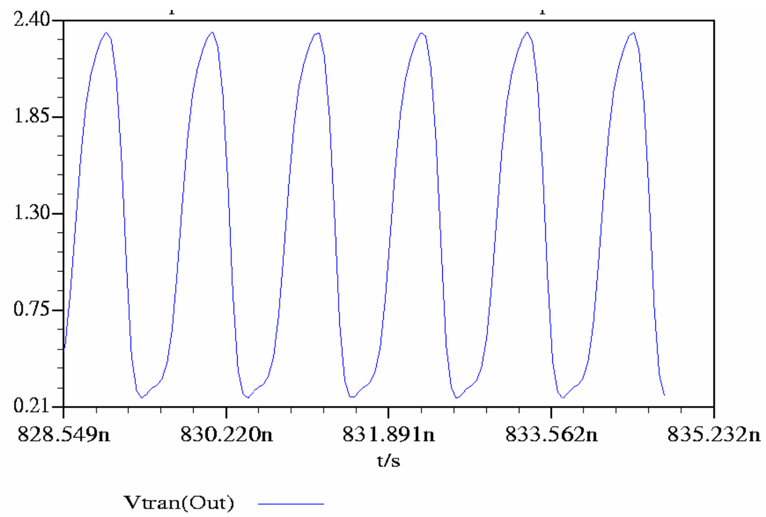


Figure 22. Simulated output waveform of the oscillator.

Due to the high Q-factor of the FBAR, start-up of the oscillator takes several hundreds of cycles. Simulated start-up time of the oscillator was approximately 800 ns (Figure 21.), frequency of oscillation 926.21 MHz and peak-to-peak amplitude 2 V (Figure 22). Current consumption of the amplifier and FBAR was 24 mA, when supply voltage Vdd was 3.3V.

A rough estimation for the standard deviation of the cycle time, or jitter, was calculated using transient noise analysis of simulator program. The simulated jitter value was 0.3 ps.

For the converter clock signal we needed in addition to a differential signal the oscillator frequency divided by 3, Figure 23. The FBAR process used did not allow straightforward manufacturing of resonators much below 900 MHz. A frequency division by 3 is not trivial to implement when having a pulse width of 50% as required by the clock generator. A method is proposed here, based on digital gates as follows. A simplified schematic diagram of the divider circuit is shown in Figure 23. Three master-slave flip-flops are connected as a ring that is initially reset to have one output state ‘1’ while the others are ‘0’. A divided signal with 50% pulse width is done with a NOR-gate taking one input from output of the first slave flip-flop instead of master as usual. Differential output is generated with an inverter while non-inverting output is delayed with an optimally sized transmission gate. A pseudo-differential cross-coupled inverter pair is inserted at the output to minimize skew between the outputs.

All circuits, excluding FBAR, were fabricated with AMS (austriamicrosystems) 0.35 µm CMOS process.

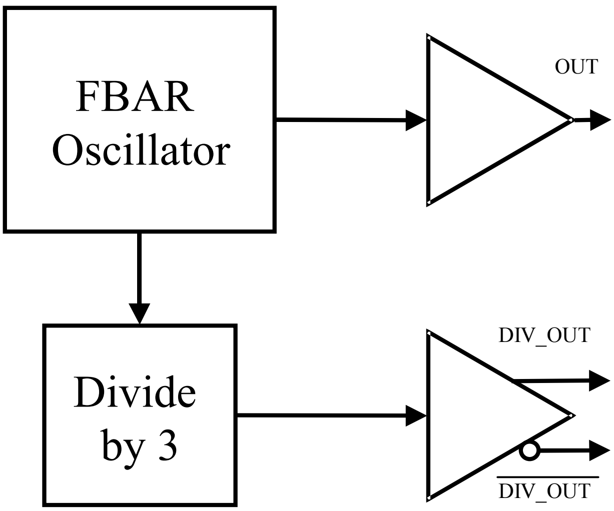


Figure 23. Block diagram of the clock circuit.

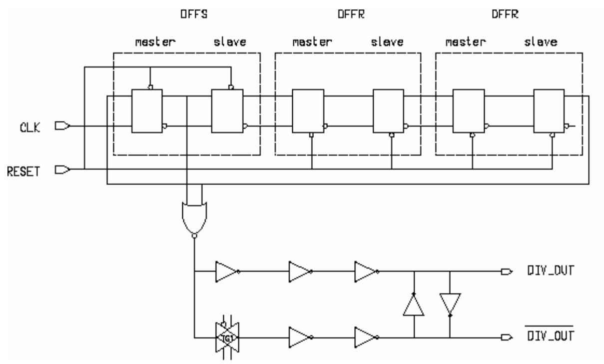


Figure 24. Divider and single-ended to differential converter.

Experimental Results

The experimental setup consisted of a CMOS-chip and a FBAR resonator wire bonded together. The measured output spectrum is shown in Figure 25 and divider output waveforms in Figure 26. The direct jitter measurement was not possible because of the resolution of the oscilloscope (minimum jitter resolution ~3 ps); only giving as a result “better than the oscilloscope resolution” (measured figure 2.7 ps). From the comparison to the measured jitter performance of a commercial clock circuit with max. 0.5 ps specified jitter (SAW oscillator type M650 provided by Integrated Circuit Systems) we could conclude that the performance of our circuit was of the same order of magnitude. We also measured the phase noise of the oscillator, Figure 25 b). Here the measurement was limited by the accuracy of the instrument, too, but we can estimate from the figure that the jitter is in the order of 0.2 ps.

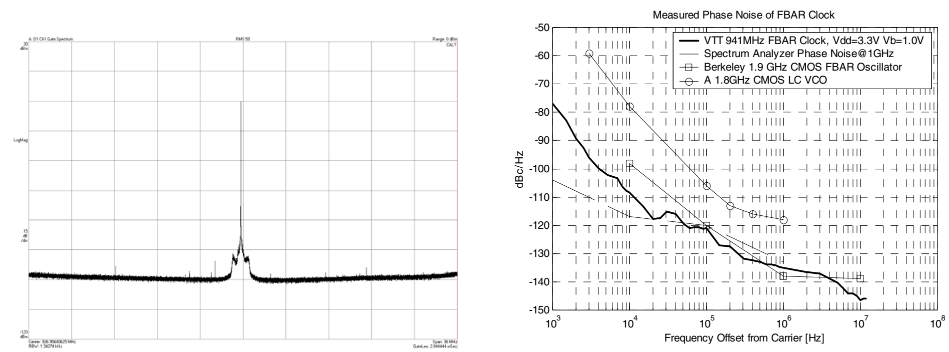


Figure 25. a) Measured output spectrum, b) Phase noise of the oscillator.

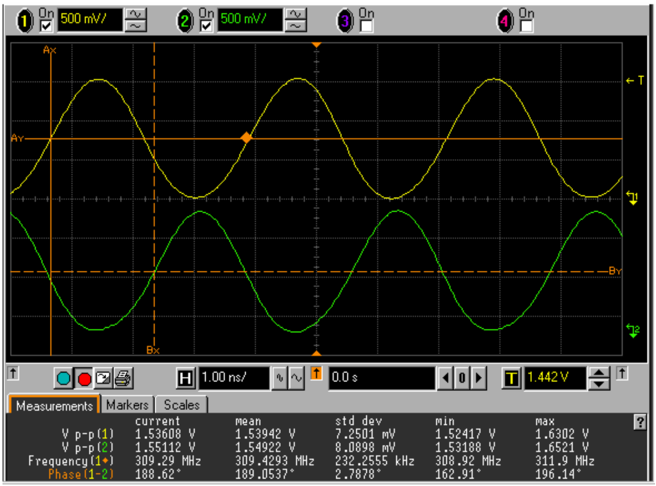


Figure 26. Measured divider output waveforms.

11.2.6 Delay Line Radio Experiments

1. Introduction

The goal of our project was to develop a new radio architecture using MEMS components. The idea was not only to replace traditional components with MEMS equivalents but to design a totally new radio architecture using MEMS components. Because we had quite short a time to develop new MEMS components, we used a MEMS delay line in our radio architecture.

2. Delay Line Architecture

The delay line radio architecture (Figure 27) consists of base station, two antennas and receiver/transmitter (called Napikka). Using a temperature sensor the Napikka sends the temperature information to the base station. The base station produces 10 μs long carrier pulses at 40 μs intervals. Napikka receives these RF pulses, add 10 μs delay and modulates the temperature information to the carrier pulses. The modulated signal will be amplified, send back to the base station receiver and demodulated. The demodulated signal is displayed using seven-segment display modules.

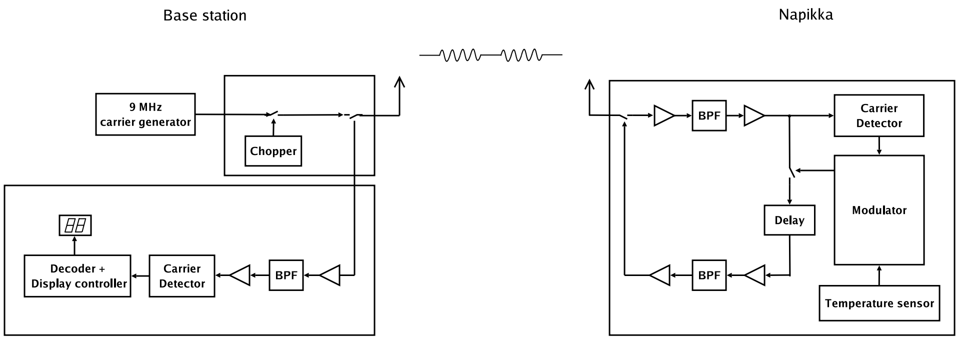


Figure 27. The block diagram of the Napikka and the base station.

Napikka does not need RF oscillators and therefore the power consumption is very low. According to the literature, most of the power in short range radio system is spent in oscillators. In this architecture the idea of the delay line is to create so long a delay that in the base station the receiver and transmitter do not operate simultaneously. Thus the base station transmitter is never active at the same time as the receiver and the high transmitter power can not block the receiver even though they are at the same frequency.

3. Napikka

The signal from the base station is received using a small loop antenna. Between the antenna and Napikka there is a switch to connect the antenna to the receiver or transmitter. The switch is followed by amplifiers and a bandpass filter. After the filter there is a diode detector and comparator to detect incoming RF pulse train. The filter is followed by amplifiers and a modulator controlled by a microprocessor. It modulates the signal from the temperature sensor onto RF pulses. The modulation is very simple on-off keying. After the modulator there are a delay line, amplifiers and a bandpass filter before the antenna relay and antenna.

4. Delay Line

Because the center frequencies of the MEMS delay line prototypes were between 9 and 11 MHz, our radio prototype worked at 9 MHz. The length of the delay line is 10 μs. The development of the MEMS delay line started at the same time as the radio architecture design. Because the MEMS delay line was not available, we needed some other delay line to test the radio architecture. Consequently, we had to use an optical fiber as a delay line. It needs a laser and photodiode to convert the 9 MHz signal to an optical signal and back to an RF signal.

5. Base Station

The base station transmitter consisted of an external RF signal generator and a switch controlled by a microprocessor. The frequency of the signal generator was 9.0 MHz and the power level was set to 0 dBm. The base station transmitted 40 μs long RF pulses at 40 μs intervals to the antenna. The signal from the antenna was connected via antenna relay to the receiver. The relay was followed by amplifiers and a bandpass filter. The signal was then amplified and detected by a diode, followed by a lowpass filter and a comparator. The microcontroller decoded the incoming temperature information and the temperature was displayed using two seven-segment display modules.

6. Antenna

The antenna element is one of the key components in any sensor system. Because the antenna of the short range transceiver must be small, we used a loop antenna, the diameter of which was 27 cm. The free space attenuation increases as the frequency increases. Considering only the free space attenuation, it is best to use as low a frequency as possible. However, in the short distance radio system at 9 MHz the dipole antenna is too long (16.7 m).

If the distance of the two 27 cm antennas is 0.5 m, the measured attenuation is 44 dB at 9 MHz. Calculating from the free space formula we get a 5 m distance at 869 MHz using a half wave dipoles with the same (44 dB) attenuation. Similarly if we get a one-meter distance using 9 MHz antennas it is equal to a 20 m distance at 869 MHz. Therefore by using small antennas, we need to use a higher transmitting frequency, 869 MHz or 2.4 GHz are good choices.

7. Conclusions

A new sensor system using delay line architecture was simulated and demonstrated. Delay line architecture facilitates a very low power consumption because no RF oscillators are needed at the sensor node. By using the 9 MHz frequency and small loop antennas the communication distance is one meter but using higher frequencies distances more than 20 meters are possible. With CMOS technology it is possible to miniaturize most of the Napikka components.

11.2.7 Packaging of RF-MEMS Devices

The packaging of RF MEMS components is very challenging because they include very delicate moving elements like membranes and in some components, very narrow gaps between the vibrating elements. The components are therefore very sensitive to mechanical stresses, contaminants and damping by air or other gaseous atmosphere. There are also challenges due to the high frequency operation of the components. That is the case for components working at frequencies of several tens of GHz. For many applications the cost of the packaging is a critical issue. Wafer level packaging is a very advantageous technology for many RF MEMS components because hermetic sealing even in vacuum can be done in a cost efficient way in a parallel process for all components of the wafer. Wafer level vacuum sealing technology has been studied at VTT Information Technology. It is however in many cases impossible to integrate all of the components of a system onto one single chip. In this case the best solution is the so called system on package (SOP) solution studied at VTT Electronics, where a high density substrate is used as a packaging platform. It serves as a platform for connections of semiconductor chips and for integrating various passive elements e.g. resistors, capacitors, inductors, filters and antennas.

When Low Temperature Co-fired Ceramics (LTCC) technology is used as a packaging solution it offers an alternative way to realize hermetic sealing of RF MEMS components into cavities made into the LTCC substrate. The sealing can be done either for the whole LTCC panel in an oven process or individually for each component. Figure 28 shows a test structure used for testing the hermetic vacuum sealing of a RF MEMS component. For this test structure a heater element was integrated into the LTCC substrate to do the sealing in a vacuum chamber. A hermetically tight sealing was made by using this technology, and the MEMS chips were assembled into the test packages. The components sealed were however not sensitive enough to test the quality of vacuum inside the sealed cavity.

The development of LTCC technology and processing has been made by studying new LTCC materials and their mutual compatibility and by studying technologies for making cavities, channels and specially formed, e.g. cylindrical structures. Interconnection technologies such as flip chip bumping and flip chip joining have also been studied. Pictures of some test structures are shown in Figure 28.

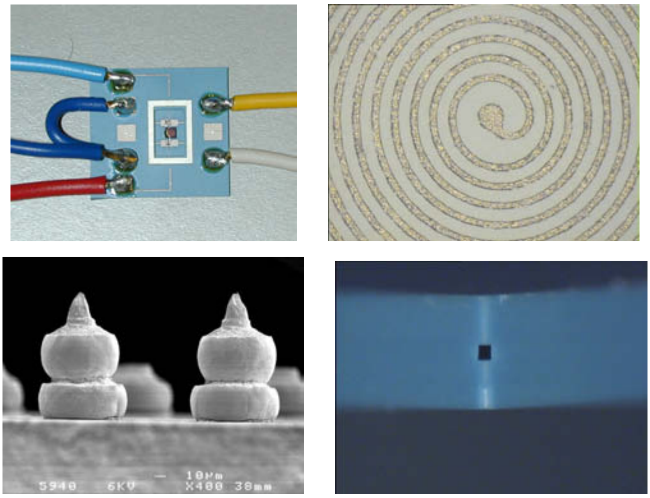


Figure 28. Test structures used for testing of vacuum sealing on LTCC, for fine line conductors and channels in LTCC and stud bump process development.

One benefit of the LTCC technology is the possibility to integrate antenna into the package. Several types of antenna test structures have been designed at VTT Information Technology and manufactured by the LTCC pilot line at VTT Electronics. LTCC technology has been used for manufacturing of various slot and patch type antennas and meander type antennas as shown in Figure 29. Trials were also made to cylindrical spiral antennas utilizing LTCC technology. Using Heralock non shrinking LTCC material it was possible to manufacture cylindrical spirals with one turn, but the technology seemed not to be mature enough to manufacture miniaturized antenna elements.

The meander type antennas shown in Figure 29 were manufactured using Dupont 943 LTCC material and the antenna had a center frequency of 921.5 MHz. Figure 29 also shows the radiation pattern of a meander antenna.

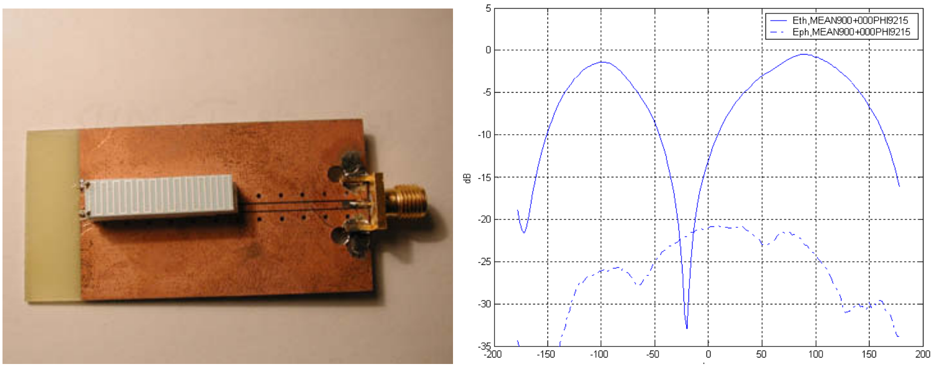


Figure 29. A meander type antenna manufactured on LTCC and its measured radiation pattern.

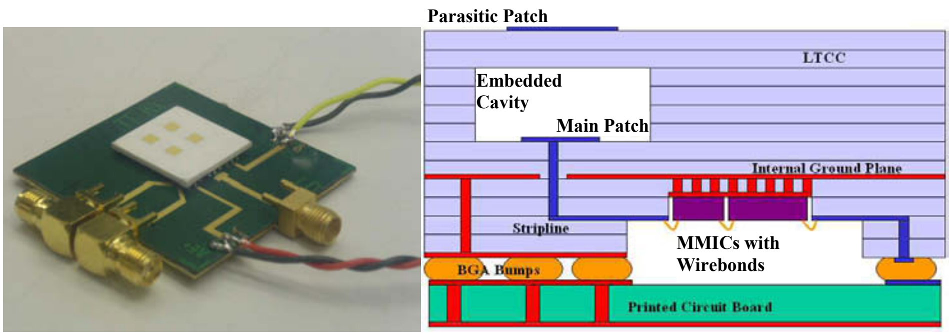


Figure 30. An LTCC front end module (as attached onto PCB) including an 2 x 2 array of embedded antenna elements as well as parasitic antenna elements on the surface of LTCC. A photograph (left) and cross section (right) [P9].

Publications

A. Journal Papers

[P1] A. T. Alastalo, V. Kaajakari, "Intermodulation in Capacitively Coupled Microelectromechanical Filters." IEEE Electron Device Letters, Vol. 26, No. 5, May 2005, pp. 289–291.

[P2] A. T Alastalo, V. Kaajakari, “Third-Order Intermodulation in Microelectromechanical Filters Coupled with Capacitive Transducers.” IEEE Journal of Microelectromechanical Systems, Vol. 15, No. 1, February 2006, pp. 141–148.

[P3] A. T. Alastalo, V. Kaajakari, “Systematic Design Approach for Capacitively Coupled Microelectromechanical Filters.” IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 53, No. 9, September 2006, pp. 1662–1670.

[P4] A. T. Alastalo, T. Mattila, H. Seppä, “Analysis of a MEMS Transmission Line.” IEEE Transactions on Microwave Theory and Techniques, Vol. 51, No. 8, August 2003, pp. 1977–1981.

[P5] A. T. Alastalo, J. Kiihamäki, H. Seppä, “Microelectromechanical delay lines with slow signal propagation.” Journal of Micromechanics and Microengineering, Vol. 16, August 2006, pp. 1854–1860.

[P6] A. Panther, A. Petosa, M. G. Stubbs, K. Kautio, “A Wideband Array of Stacked Patch Antennas Using Embedded Air Cavities in LTCC.” IEEE Microwave and Wireless Components Letters, Vol. 15, No. 12, December 2005, pp. 916–918.

[P7] Rantala, A., Martins, D. G., Åberg, M. “A DLL clock generator for a high-speed A/D converter with 1 ps jitter and skew calibrator with1 ps accuracy in 0.35 um CMOS.” Analog Integrated Circuits and Signal Processing, Vol. 50, January 2007, pp. 69–79.

[P8] Åberg, M., Ylimaula, M., Ylilammi, M., Pensala, T., Rantala, A. “A Low Noise 0.9 GHz FBAR Clock.” Analog Integrated Circuits and Signal Processing, Vol. 50, January 2007, pp. 29–37.

[P9] Stubbs, M. G., Kautio K. “A Compact 25–30 GHz LTCC Front End with Integrated Antenna Array.” (A manuscript to be published.)

B. Conference Papers

[P10] A. T. Alastalo, T. Mattila, H. Seppä, J. Dekker, "Micromechanical slow-velocity delay lines." Proc. 33rd European Microwave Conference, Munich, Germany, 6–10 October, 2003. Pp. 967–970.

[P11] A. T. Alastalo, M. Koskenvuori, H. Seppä, J. Dekker, "A Micro-mechanical Resonating RF Mixer." In: 34th European Microwave Conference, 2004. Pp. 1297–1300.

[P12] A. T. Alastalo, V. Kaajakari, "Designing Capacitively Coupled Micro-electromechanical Filters." In: IEEE International Ultrasonics Symposium, 2005. Pp. 1588–1591.

[P13] V. Kaajakari, A. T. Alastalo, T. Mattila, "Electrostatic transducers for micromechanical resonators: free space and solid dielectric." IEEE Trans. Ultrason., Ferroelect., Freq. Contr., 2006, (In press.)

[P14] J. Dekker, A. Alastalo, "Wafer Level Encapsulation of Soi Narrow Gap Resonators." 2nd International Workshop on Wafer Bonding for mems technologies, Halle, Germany, April 2006

[P15] A. Rantala, D. G. Martins, M. Åberg, "A DLL based clock generator for a high speed time-interleaved A/D-converter." In: Proc. of IEEE Norchip 2003 Conference, Riga, Latvia November 2003.

[P16] A. Rantala, D. G. Martins, M. Åberg, "A DLL clock generator for a high-speed A/D converter with 1 ps jitter and skew calibrator with1 ps accuracy in 0.35 um CMOS." In: Proceedings of the 23rd Norchip Conference. Oulu, 21–22 Nov. 2005. TechnoData A/S. Faaborg, DK (2005). Pp. 114–117.

[P17] M. Åberg, M. Ylimaula, M. Ylilammi, T. Pensala, A. Rantala, Arto. "Low noise 0.9 GHz FBAR clock." In: Proceedings of the 23rd Norchip Conference. Oulu, 21–22 Nov. 2005. TechnoData A/S. Faaborg, DK (2005). Pp. 257–260.

[P18] A. Rantala, D. G. Martins, M. Åberg, "A DLL clock generator for a high-speed A/D converter with 1 ps jitter and skew calibrator with1 ps accuracy in 0.35 um CMOS." In: Proceedings of the 23rd Norchip Conference. Oulu, 21–22 Nov. 2005. TechnoData A/S. Faaborg, DK (2005). Pp. 114–117.

[P19] M. Sipilä, P. Raatikainen, M. Åberg, I. Norros, L. Norros, A. Mämmelä, T. Frantti, J. Holmberg, H. Ailisto, A. Lattunen, J. Latvakoski, S. Horsmanheimo, "Topics in 4G networks and hardware" In: Proceedings of the International Workshop on Optical and Electronic Device Technology for Access Network." Aalborg, 2005. Aalborg University, Keio University. Tokyo, Japan (2005). Pp. 5–10.

C. Other Publications

[P20] A. Oja, A. Alastalo, J. Knuuttila, A. V. Räisänen, I. Tittonen, A. Kallio, M. M. Salomaa, "MEMS-radio yhdelle piisirulle." Prosessori-lehden erikoisjulkaisu, marraskuu 2003, pp. 54–55.

D. Patents and Invention Announcements

[P21] H. Seppä, A. Alastalo, "Oskillaattorikytkentä ja menetelmä sekä mikromekaaninen vaihemonikäsitteinen kaistanpäästösuodatin", Patenttinumero: FI 114668 (30.11.2004), (VTT).

[P22] H. Seppä, A. Alastalo, "Kytkentä ja menetelmä lähetinvastaanottimia varten sekä mikromekaaninen viivelinja", Patenttinumero: FI 20031427 (3.4.2004), (VTT).

[P23] S. Tallqvist, "Rolled up Cu spiral antenna" Keksintöilmoitus, VTT 2003.

E. Theses

[P24] A. Alastalo, "Microelectromechanical Resonator-Based Components for Wireless Communications: Filters and Transmission Lines", Dr. Tech. Dissertation, Helsinki University of Technology, to be defended in dissertation on November 1, 2006.

References

[R1] A. B. Carlson, P. B. Crilly, J. C. Rutledge, "Communication Systems." McGraw-Hill, 4th edition, 2002.

[R2] V. Kaajakari, T. Mattila, A. Oja, J. Kiihamäki, H. Seppä, "Square-Extensional Mode Single-Crystal Silicon Micromechanical Resonator for Low Phase Noise Oscillator Applications." IEEE Electron Device Lett., Vol. 25, No. 4, 2004, pp. 173–175.

[R3] P. Gammel, G. Fischer, J. Bouchaud, "RF MEMS and NEMS Technology, Devices, and Applications." Bell Labs Technical Journal, Vol. 10, No. 3, 2005, pp. 29–59.

[R4] S. Lucyszyn, "Review of radio frequency microelectromechanical systems technology." IEE Proc.-Sci. Meas. Technol., Vol. 151, No. 2, 2004, pp. 93–103.

[R5] T. Vähä-Heikkilä, "MEMS tuning and matching circuits, and millimeter wave on-wafer measurements." D.Sc. thesis, Helsinki University of Technology, VTT Publications 596, Espoo, 2006. 86 p. + app. 82 p. ISBN 951-38-6704-8; 951-38-6705-6. http://virtual.vtt.fi/inf/pdf/publications/2006/P596.pdf.

[R6] C. T.-C. Nguyen, "Frequency-Selective MEMS for Miniaturized Low-Power Communication Devices." IEEE Trans. Microwave Theory Tech., Vol. 47, No. 8, 1999, pp. 1486–1503.

[R7] J. J. Yao, "RF MEMS from a device perspective." J. Micromech. Microeng., Vol. 10, 2000, pp. R9–R38.

[R8] V. K. Varadan, K. J. Vinoy, K. A. Jose, "RF MEMS and Their Applications." John Wiley & Sons Ltd, 2003.

[R9] G. M. Rebeiz, "RF MEMS, Theory, Design, and Technology." John Wiley & Sons, Inc., 2003.

[R10] H. Seppä, J. Kyynäräinen, A. Oja, "Microelectromechanical Systems in Electrical Metrology." IEEE Trans. Instrum. Meas., Vol. 50, No. 2, 2001, pp. 440–444.

[R11] A. Alastalo, J. Kyynäräinen, H. Seppä, A. Kärkkäinen, N. Pesonen, M. Lahdes, T. Vähä-Heikkilä, P. Pekko, J. Dekker, "Wideband microwave power sensor based on MEMS technology." In: Conference on Precision Electromagnetic Measurements (CPEM 2004), 2004. Pp. 115–116.

[R12] L. Fernández, J. Sesé, R. Wiegerink, J. Flokstra, H. Jansen, M. Elwenspoek, "Radio Frequency Power Sensor Based on MEMS Technology with Ultra Low Losses." In: IEEE International Conference on Micro Electro Mechanical Systems, (MEMS 2005), 2005. pp. 191–194.

[R13] A.-C. Wong, C. T.-C. Nguyen, "Micromechanical Mixer-Filters ("Mixlers")." J. Microelectromech. Syst., Vol. 13, No. 1, 2004, pp. 100–112.

[R14] C. T.-C. Nguyen, "Method and apparatus for generating a signal having at least one desired output frequency utilizing a bank of vibrating micromechanical devices." United States Patent 6,577,040, 2003.

[R15] B. P. Otis, Y. H. Chee, R. Lu, N. M. Pletcher, J. M. Rabaey, "An Ultra-Low Power MEMS-Based Two-Channel Transceiver for Wireless Sensor Networks." In: Symposium On VLSI Circuits, 2004. Pp. 20–23.

[R16] R. Navid, J. R. Clark, M. Demirci, C. T.-C. Nguyen, "Third-order intermodulation distortion in capacitively-driven cc-beam micro-mechanical resonators." In: IEEE International Conference on Micro Electro Mechanical Systems, (MEMS 2001), 2001. Pp. 228–231.

[R17] R. B. Reichenbach, M. Zalalutdinov, K. L. Aubin, R. Rand, B. H. Houston, J. M. Parpia, H. G. Craighead, "Third-Order Intermodulation in a Micromechanical Thermal Mixer." J. Microelectromech. Syst., Vol. 14, No. 6, 2005, pp. 1244–1252.

[R18] A. Rantala, D. Gomes Martins, M. Åberg “A DLL based clock generator for a high speed time-interleaved A/D-converter.” In: Proc. of IEEE Norchip 2003 Conference, Riga, Latvia, November 2003.

12. Fast Analog and Digital Signal Processing in Wireless Communications (NASK4G)

Aarne Mämmelä, Jan Holmberg, Atso Hekkala, Pasi Hyttinen,
Vasily Kondratyev, Henna Paaso, Yang Qu, Arto Rantala,
Juha-Pekka Soininen, Antti Tanskanen, Leo Vaskelainen and Juha Volotinen

12.1 Introduction

The processing in analog and digital domains and their co-operation is a critical problem in wireless communications where the efficiency of the system in terms of energy consumption is important. In order to better handle the challenges of the future mobile applications, exhaustive system level considerations must be carried out.

A common trend is to digitize physical signals as early as possible and perform most of the signal processing in the digital domain. Often this is feasible and effective procedure, but can occasionally lead to impractical implementations if one or several wideband or high frequency signals are involved. An example of such a case is a smart array antenna where several simultaneous signals must be processed. Analog signal processing and computing circuits provide real time true parallel processing with continuous signal values. Analog circuits are power efficient when compared to digital counterparts, but usually are less accurate due to process and environmental sensitivity. In addition the technology being the same, analog operations can be performed always in a higher speed than their digital counterparts.

There are a wide range of versatile circuits for most analog signal processing tasks. However, two critical functions, delay and memory, are difficult to implement. In this work some feasible solutions are investigated and developed.

Smart antenna concepts are becoming important as the data rates and spectral efficiency requirements increase in fourth generation (4G) wireless connections. One possible way to improve the spectral efficiency of applications is to use antenna arrays with narrow and adaptive beams. Their systematic design poses a difficult mathematical problem. In this study a novel optimisation method for beamforming of smart antennas is presented.

Another important problem is modeling and compensation of nonlinear effects of the analog parts of the transmitter and receiver in the future 4G mobile cellular systems. Compensation may be analog or digital. We have concentrated on the nonlinearity of high power amplifiers in the transmitter. The nonlinearity will be an important design challenge in future systems where the signals (for example, quadrature amplitude modulation (QAM) and orthogonal frequency division multiplexing (OFDM) signals) do not any more have a constant envelope, they are very sensitive to the various nonideal effects, and the signals have a large bandwidth.

Finally, architecture, design, and evaluation of software and hardware communication platforms were studied. The plan was to develop methods for evaluation and design of digital baseband processing architectures for future communication platforms. Reconfigurability, on-chip networks, and multiprocessor architectures were key technological topics.

12.2 Analog Signal Processing

At the same time as more complex signal processing is needed in the new emerging wireless applications the price of the user terminals should be falling. Although an ever increasing share of the terminal will be digital the performance of the analog front end becomes more critical. Especially more effective antenna systems should be developed. In order to facilitate sophisticated front end implementations we have chosen three main research topics, which are true time delay circuits, floating gate tuning structures and smart antenna beam synthesis.

In order to integrate wide band true time delay circuits for low microwave frequencies in a small volume active implementations are needed. In many cases some kind of tuning of analog parts is needed to adapt the circuit into current situation. Traditionally tuning voltages or currents are set by trimming potentiometers, but they are not well suited either for integration or mass production. Nowadays usually digitally controlled potentiometer chips are used instead. This implies naturally quantized tuning values which could be avoided by fully analog implementation. Floating gate structures might offer a solution to this problem.

12.2.1 True Time Delay Cell

In high frequencies distributed structures can be used, but in the range below 10 GHz they become too space consuming and integration level will be low. A common approach is to implement the delay block using passive elements to provide lumped element approximation for transmission line. A delay can also be implemented by using active components. For example bipolar junction transistors (BJT) have process dependent (both geometry and doping) emitter-to-collector transit time that can be utilized in delay circuits. The BJTs are also very suitable for current mode signal processing.

The developed delay blocks utilize current mode signalling and translinear principles. A promising delay circuit is based on an improved Wilson current mirror circuit. It provides wideband characteristics (Figure 1) and robust design to cope with process variations. The size of the five cell circuit is only about 60 µm x 150 µm. Another delay circuit topology is based on emitter coupled differential pair with translinear input stage that provides improved performance. The tuning and biasing of the differential delay is more robust and simpler than Wilson’s current mirror, but differential pair has a somewhat worse frequency response.

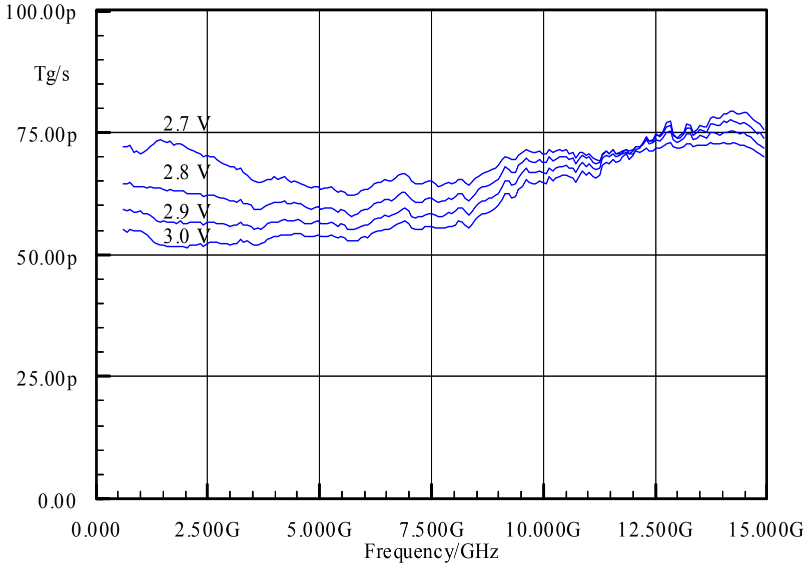


Figure 1. Measured group delay curves of the true time delay circuit consisting of five unit cells in cascade with the control voltage as a parameter.

Analog finite impulse response (AFIR) filter is a versatile programmable signal processing block that could be utilized in many different applications. It can be used for example in adaptive filtering, signal generation, and linearization of memory effects in power amplifiers. The AFIR consists of a chain of analog delay blocks, taps with tunable gain, and a summing element. An implementation of a four tap AFIR is shown in Figure 2.

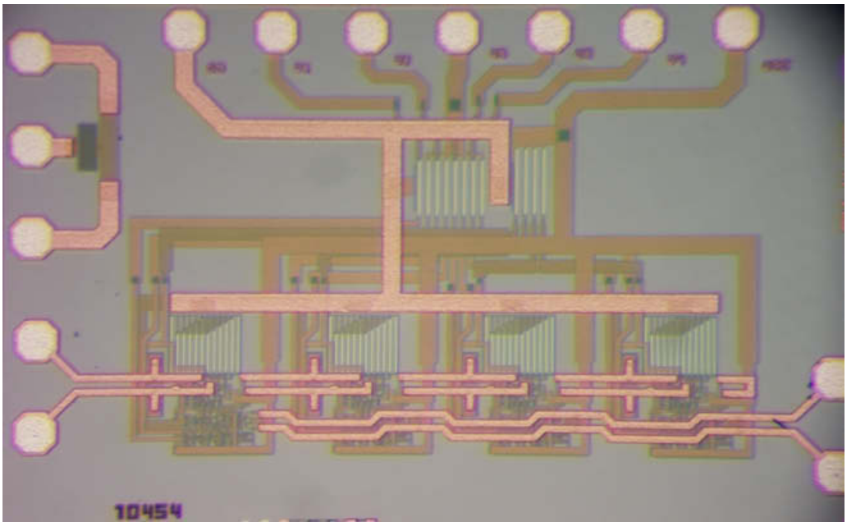


Figure 2. Microphotograph of a 4-tap analog FIR filter.

12.2.2 Floating Gate Analog Memory Cell

In many integrated circuits some data, digital or analog, has to be stored for a long time, even without a power supply. In many cases the flexibility of the system can be enhanced if the values can be changed afterwards. In addition to digital values it is sometimes required to store an analog parameter. Then it can be advantageous to store bias voltages or calibration parameters as analog values. The proposed method to form non-volatile memory is based on a floating gate structure studied earlier [P6].

A floating gate can be charged and discharged to form an electrically erasable programmable read-only memory (EEPROM) in CMOS either using Fowler-Nordheim-tunnelling or Channel Hot-Electron Injection. Many topologies have been investigated and found that the memory structure which fulfils the requirements in the best manner should have following features [P6]:

- Fowler-Nordheim-tunneling based programming (low current)
- Tunneling through thin oxide (charging and discharging)
- Separated thick oxide transistor for charge measurement.

Nonvolatile memory can be formed by having a floating node, which preserves the applied charge. In the simplest form EEPROM is a MOS-transistor having floating gate. The charge stored in the gate can be altered, which makes a shift to the threshold voltage of the transistor. In the digital EEPROM the shift of the threshold voltage is detected as a logic state i.e. with a sense amplifier. When an analog value (i.e., controlled amount of charge) is stored into the EEPROM, the floating gate transistor can be utilized e.g. as a multistate or analog programmable current source. The control signals for the EEPROM can be connected capacitively to the floating gate.

A schematic diagram of the proposed EEPROM memory cell is shown in Figure 3. The charge tunnels through the gate of the transistor when either VT or VG are connected to high voltage while the other node is at a lower potential. Since the insulators of the memory read transistor at the amplifier input and polysilicon capacitor are much thicker, the floating gate (FG) of the transistor is the only tunnelling path. In a normal operational mode of analog memory the negative input is connected to the output of the amplifier. Thus a voltage follower circuit having a nonvolatile, programmable offset is obtained.

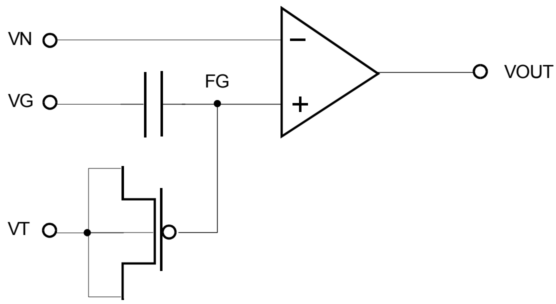


Figure 3. A schematic diagram of the proposed EEPROM analog memory cell.

The operation of the analog EEPROM was characterized by generating accurately controlled voltage pulses to VT and VG inputs while the output voltage the operational amplifier was measured. An iterative algorithm controlled the charging until the required level and accuracy was obtained. After the programming an offset is observed in the voltage transfer characteristics of the voltage follower. It turned out that the analog memory works in quite a linear manner within the required voltage range.

A typical application could be in voltage controlled oscillators (VCO). In communication circuits the most widely used topology is an LC oscillator having controllability over frequency with a tunable capacitor. This topology can be completely integrated by using a planar inductor. The major drawbacks of an integrated LC oscillator are low tuning range (due to the low tuning range of the varactor) and oscillation frequency variation (due to the IC processing tolerance). As a result there is a trade-off between tuning range and control sensitivity. In this study a method is proposed to utilize an analog nonvolatile memory (EEPROM) to calibrate the LC oscillator frequency.

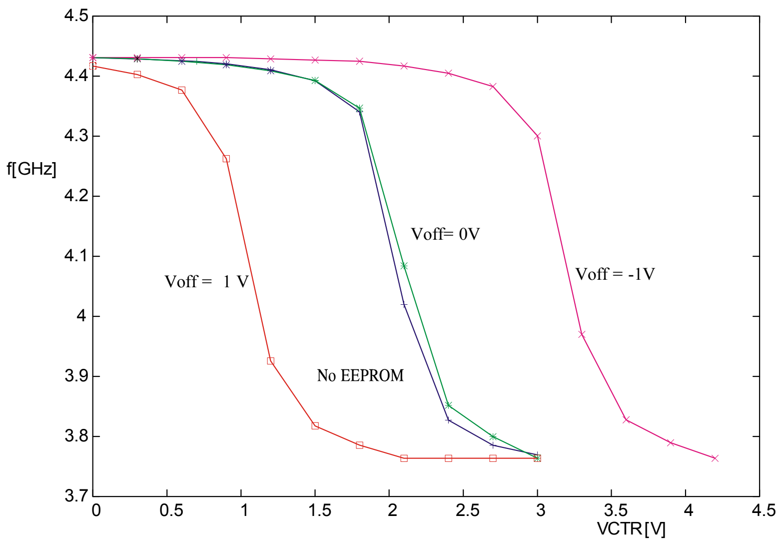


Figure 4. The measured output frequency of the VCO as a function of the control voltage V_{CTR} at three offset values V_{off}. The curve labelled "No EEPROM" is measured in the case of direct control of the VCO.

The frequency calibration performance of the VCO was tested using a programmed analog memory cell. The developed memory cell was connected to a VCO [P7]. The control transfer characteristics were measured using three programmed offset values V_{off} and results were compared to values obtained from circuit which do not include the EEPROM. The measured results of the cases (depicted in Figure 4) show that the proposed idea is functional and the oscillation frequency can be adjusted nonvolatively.

12.2.3 Array Antenna Beam Optimisation in Smart Antennas

The synthesis problem of an arbitrary array antenna can be seen as a general optimisation problem, where optimal complex element excitation values are sought. To complete this kind of optimisation, global optimisation methods are usually used, such as genetic algorithms, simulated annealing or particle swarm optimisation. Global optimisation methods are very flexible to use, but they need a large amount of calculation of directivity patterns.

Least-squares optimisation in its basic form gives directly a solution for the unknown excitation values. Unfortunately this solution is for some guessed phase values of the goal function, which does not necessarily give optimum fit for the synthesised directivity function amplitude. For finding the best amplitude fit some iterative method must be used.

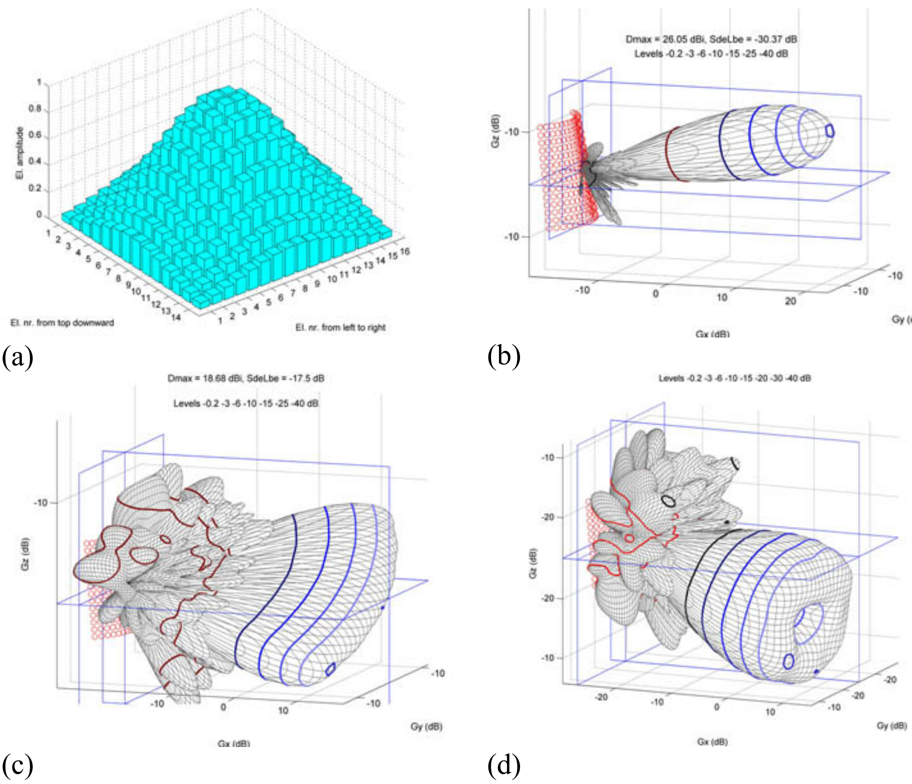


Figure 5. Pencil-beam (b), wide fan-beam (c), and a circular beam with a null (d) synthesised using only phase optimisation and the same Gaussian amplitude distribution (a). There are 224 antenna elements on an elliptical cone.

Least-squares optimisation is also not directly suitable for phase synthesis, which is a nonlinear optimising problem. In smart conformal array antennas a phase synthesis method is important because we want to do the beam control by using only the phase control and some constant pre-defined amplitude distribution. In an optimal solution we can use the same amplitude distribution and change the beam direction and the beam shape by phase control.

In this work the minimized function (least-mean square (LMS) error) of the constrained least-squares optimisation problem is studied. It is found that the minimised function for phases of the goal function and possible constraint values has a nonlinear, but very simple form. In addition in this minimised function only the nonzero values must be considered, and the number of nonzero goal values is usually notably smaller than the total number of directions used in the synthesis. Therefore the optimisation of these phase values is a very fast process. Fast optimisation technique for the phases of constraint values allows this method to be used also for the phase synthesis.

Basically the least-squares optimisation controls only the general level of side-lobes, not the level of individual peaks. The optimisation error in different directions can be controlled by the selection of weights used in the minimized function. Iterative control of individual error values can be achieved by correction of the weight values according to the previous synthesis error values.

The details of the new synthesis method of conformal array antennas are presented in [P8]. Very large phase synthesis problems can be solved by using this method, as presented in [P9]. Example results of a large conformal phase-synthesis problem are presented in Figure 5.

12.3 Algorithms

Linear amplification is required when the signal contains both amplitude and phase modulation. If the nonlinearity has memory, it may distort even constant amplitude signals. Nonlinear distortion can be thought of as the creation of undesired signal energy at frequencies not contained in the original signal. A solution for reducing the adjacent channel interference is to linearize the power amplifier (PA) used in the transmitter. The second motivation for the linearization is the improved efficiency of the amplifiers operating in nonlinear modes. The predistortion algorithms require reliable models of the power amplifiers. Therefore we simulated a practical power amplifier model. In predistortion we considered two approaches, one based on indirect learning architecture and the other based on direct learning architecture. These two approaches were compared.

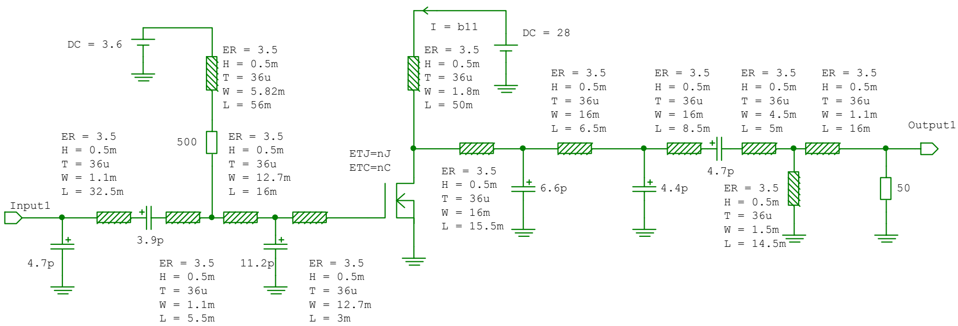


Figure 6. Amplifier simulation model.

12.3.1 Modelling of Power Amplifiers

In this work, some transmitter architectures for linear amplification using nonlinear amplifier blocks are introduced on the basis of a literature review. Methods to linearize linear amplifier blocks even further are discussed and classified. Of these methods, the predistortion is highlighted. For the successful design and simulation of a linearized amplifier, an amplifier model including various nonlinear effects with and without memory is required. For this purpose, a practical 75-W solid-state amplifier is designed and simulated (Figure 6). A simple thermal model for the active device is included in the modelling in order to capture the thermal memory effects. Simulation results on AM-AM, AM-PM conversions and amplifier memory effects are presented. Finally, a third order Volterra model of the simulated amplifier is extracted. All the simulations are performed with APLAC simulation software.

The experimental verification of the modelling work was beyond the scope of the project. However, the results, especially the phase response for the intermodulation products (Figure 7), were in line with what can be expected from this type of amplifier.

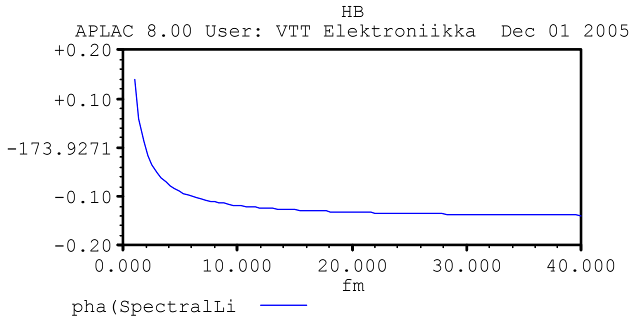


Figure 7. The phase response [degrees] of the third order intermodulation product as a function of tone separation [Hz]. The phase shift due to the thermal memory is clearly visible on lower tone separations.

12.3.2 Predistortion Algorithms

The purpose of this work is to present the time-discrete modeling and simulation of nonlinear effects and their digital compensation. In general, the physical channel itself is a major source of noise and disturbances. However, the analog parts in communication systems cause additional distortions. Practical mixers are nonlinear devices, and their imperfections include phase noise, I/Q imbalance and DC offset. All amplifiers are nonlinear if driven hard enough; they all exhibit saturation effects. Additionally, QAM/OFDM methods will be more commonly used in the future. These modulation methods have nonconstant envelopes. It can be also assumed that nonlinearities have memory due to wide bandwidth, which implies frequency-dependency to the model. However, it is not guaranteed that the nonlinear PA with memory can amplify a constant envelope without introducing significant distortion. Digital predistorters (PD) with memory can compensate these nonlinearity effects. Several memoryless predistorters are presented in the literature, but predistorters with memory are much less known.

In the research, the main attention was given to the two memory polynomial predistorters: direct ([R2]) and indirect ([R1], [R3]) learning architectures. Memory polynomial model can be expressed as in [R1]

$$y_n = \sum_{k=1}^{K} \sum_{q=0}^{Q} c_{kq} z(n-q) \left| z(n-q) \right|^{k-1}$$

where K is the order of the nonlinearity, Q is the memory length, c_{kq} is a coefficient of the nonlinear model, $z(n)$ is the input signal of the nonlinear model. Direct and indirect learning architectures are special cases of the self-tuning control. These models are illustrated in Figure 8. Direct learning means that the estimate of the input-output relation of the power amplifier is updated for the characterization of the PA, and the predistortion is obtained directly by "pre-inverting" the PA characteristics. Indirect learning means that the postdistorter first derives a postinverse of the nonlinear model without any predistorter and then the postdistorter is used as a predistorter, and the postdistorter is removed. Inversion of a nonlinear system may not be possible: not all nonlinear systems possess an inverse, and many systems can be inverted only for a restricted range of input. Invertible nonlinear model transformation must be a bijection, i.e., there is one-to-one correspondence between the input and the output. In this case the inverse is also a function.

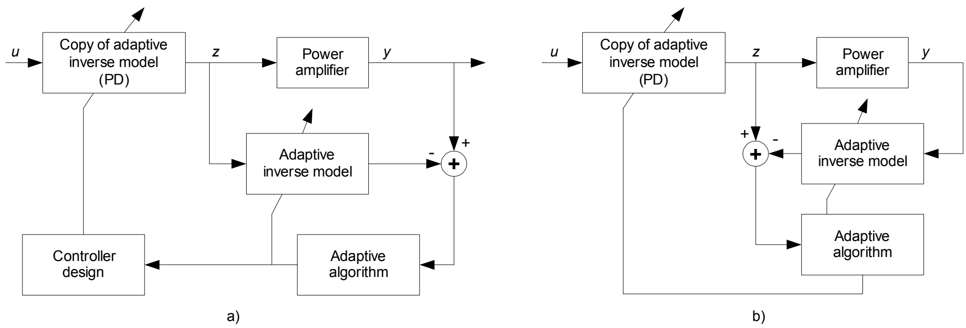


Figure 8. Direct (a) and indirect (a) architectures for the predistorter.

First, the AM/AM and AM/PM characteristics were studied at different frequencies in linear and nonlinear cases by using Matlab. Simulation results showed that both AM/AM and AM/PM characteristics are independent of frequency in memoryless linear cases. In nonlinear memory cases, we noted that memory polynomial PA is frequency dependent and the nonlinear characteristics are clearly visible. In our simulations we noted that the memory polynomial model has convergence problems at large amplitudes. This is typical for Volterra based models.

Finally, the nonlinearities of the memory polynomial PA are linearized by using a memory polynomial predistorter. We used two memory polynomial predistorters: direct ([R2]) and indirect ([R1]) learning architectures. Example results of the AM/AM and AM/PM characteristics for the PD and PA entity are presented in Figures 9 and 10. Figure 9 shows the result for the direct learning predistorter and Figure 10 the result for the indirect learning predistorter. In both the cases nonlinearities can be compensated if the amplitude of the input signal is lower than one. However, if the amplitude of the input signal is higher than one, the PA nonlinearities of the PA cannot be compensated. The predistorter also tries to compensate the saturation area. It compensates harder this area if the amplitudes of the measured signal are higher. Consequently, we can conclude from this that the result of the compensation depends also on the amplitude distribution, not only on the frequency. The other simulation results show that a direct model achieves a better performance in almost all cases. However, the results also show that a direct learning model has a stability problem. The direct learning model is time-variant because each time sample has its own coefficients of the predistorter. The indirect learning model is time-invariant. Therefore, the coefficients of the predistorter do not depend on time. This is one reason for the difference in the performance.

In the indirect learning architecture ([R1]) the problem of commutability must be considered ([P2]). We pointed out that linearity is not a necessary nor sufficient requirement for the commutability since matrix multiplication is not in general commutative although scalar multiplication is commutative. For linear time-invariant systems that obey the superposition principle we can use the product of two transfer functions, which is essentially a set of scalar multiplications, and therefore the blocks commute. On the other hand, some blocks are represented by a matrix, for example linear time-variant blocks and blocks describing I/Q imbalance. Such blocks do not in general commute unless the matrices have some special properties. Furthermore, nonlinear systems are not in general commutative. The commutability is not valid unless there is a special reason for that. Typical examples for commutability include (1) the systems are combined through a commutable operation, or (2) one of the systems is the inverse of the other system.

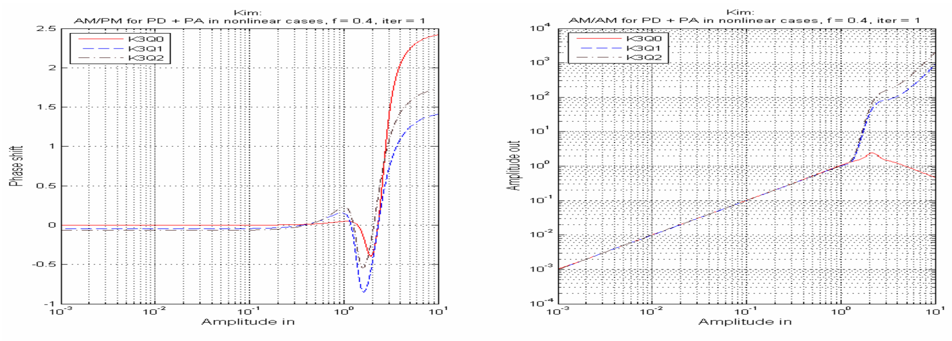


Figure 9. AM/AM and AM/PM characteristics for the PA and the direct learning predistorter entity after the first iteration, f = 0.4.

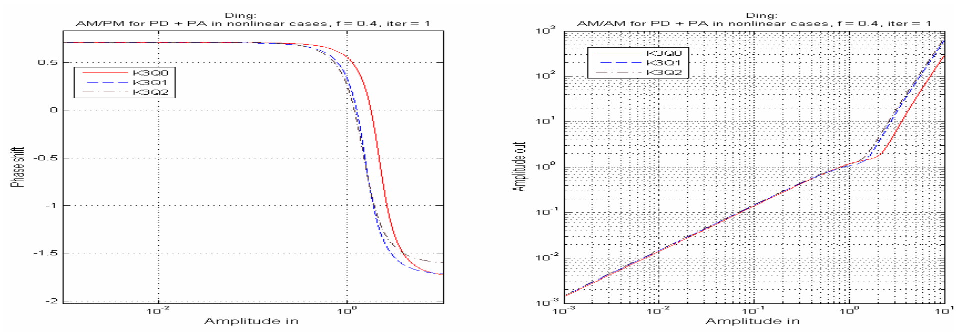


Figure 10. AM/AM and AM/PM characteristics for the PA and the indirect learning predistorter entity after the first iteration, f = 0.4.

12.4 Architectures

Design effort in future communication platforms is moving from hardware towards software solutions. Time to market and technology development are the main driver and enabler for this trend, respectively. Software design cycles are cheap when compared to hardware design cycles.

A gap between hardware and software design methodologies and performance is diminishing. Reconfigurable logic is an important design alterative in design of system-on-chip (SoC) because it can achieve better performance than software implementation while maintaining a higher degree of flexibility than fixed hardware. High silicon reusability can also be achieved through run-time reconfiguration (RTR). The RTR means that the circuit or a part of it can be reconfigured while the rest of the system is running. However, it results in configuration overhead, which can largely degrade the system performance. Therefore, approaches to model the configuration overhead for system-level performance evaluation and approaches to reduce the effect of the configuration overhead are needed in order to allow such devices can be efficiently utilized.

12.4.1 System Level Management of the Reconfiguration Overhead

Reconfigurable logic is more favorable than software implementation and fixed-hardware implementation when both the performance and the flexibility are concerns. High silicon reusability can also be achieved through run-time reconfiguration (RTR). Such devices are referred to as dynamically reconfigurable hardware (DRHW). The RTR allows tasks that are non-overlapped either in time domain or in space domain can be mapped onto the same DRHW to increase the design efficiency, but it results in the configuration overhead, e.g. configuration latency and power consumption, which can largely degrade the system performance. At the system level, methods to estimate and model such configuration latency are needed in order to accurately evaluate the system performance when DRHW is used in design. In addition, novel devices and approaches to reduce the effect of the configuration latency are needed in order to increase the efficiency of using DRHW.

In this work, we present a SystemC-based modeling approach ([P1], [P3]) for DRHW. Our approach is not intended to be a universal solution to support the design of any reconfigurable system. Instead, we focus on using reconfigurable components mainly as co-processors in SoCs. The objective of the SystemC based approach is to provide a mechanism that allows designers to easily test the effects of implementing some components in the dynamically reconfigurable hardware. The main contributions are:

- Estimation approach for design space exploration: Tasks that are to be implemented are modeled in C code. A set of high-level synthesis steps, including as-soon-as-possible scheduling, as-late-as-possible scheduling, force-directed scheduling and resource allocation, are performed to estimate hardware resources required for these tasks. The current estimator targets a Virtex2-like embedded field programmable gate array (FPGA) in which main resources are lookup-tables (LUT's) and multipliers. The estimated number of resources can help designers to decide how to map the tasks onto DRHW.

- Model reconfiguration overhead: The configuration latency is modeled as pure delay. A parameterized SystemC module that models the behavior of RTR process is created. In simulation, it can automatically capture the reconfiguration request and insert the configuration overhead, so that designers can easily evaluate the performance impact of using DRHW. To reduce the coding efforts and increase the design reusability, a code transformation tool to automatically create a new SystemC model from an existing system model is developed.

To reduce the effect of the configuration latency, we proposed a parallel configuration model ([P4]) and three static scheduling techniques ([P4], [P5]). The model consists of a number of continuously connected homogeneous tiles, and each tile consists of the circuit and its own configuration-SRAM that controls the circuit. A task that requires m tiles of resources can use any set of m connected tiles. A crossbar connection is used to connect the configuration-SRAMs of the tiles to a number of parallel configuration controllers. The crossbar ensures that any configuration-SRAM can be accessed by any controller but only one at a time. Thus, reconfigurations can be performed in parallel on different tiles. The memory that holds the configuration data of all tasks should have multiple read ports or consist of a number of separate memories that allow parallel access, so different controllers can read data at the same time (Figure 11). With a set of randomly generated task graphs, the experiment results reveal that in the average using multiple controllers can reduce the configuration overheads by 21%. Compared to best cases of using multiple tiles with a single controller, an additional 40% speedup can be achieved using multiple controllers.

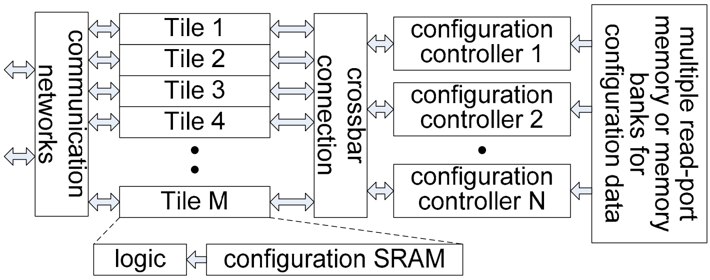


Figure 11. Parallel configuration model.

Task scheduling on such parallel configuration models is an NP-hard problem, as the task scheduling, allocation, configuration prefetching all need to be considered in the optimization process. Three different techniques are studied to solve the problem. The first is a list-based heuristic scheduler ([P4]). The basic idea is to assign tasks certain priorities and schedule them based their priorities. The priority of a task depends on the current scheduling status and structure of the directed-acyclic graph that represents the set of tasks. The second scheduler is based on constraint programming (CP, in [P5]). An application is developed that can automatically generate the constraint model for a given task graph and a given device model. The SICStus finite domain solver ([R4]) is used to solve the model. The last scheduler used genetic algorithm (GA), a guided random search strategy. The case studies show that the CP based approach is suitable for small-scale applications only, as the required computer runtime increases exponentially with the number of tasks. The list-based approach is the fastest approach, requiring a few milliseconds in the average, but the average error is about 13% in large task graphs. The GA approach in average requires a few seconds, but the average error is below 4.5%.

12.5 Conclusions

All three main research topics of analog parts have been fairly successful. However, more research is needed to pin down the actual limits of their performance. Several potential improvements are found and next evolution versions are in process. So far it was demonstrated that active true time delay circuits can be implemented on a very small chip area, so that their integration is possible. Floating gate analog memory cell was developed, implemented and characterized and could be used in various tuning applications. As a pure analog circuit it offers superior resolution compared to the digital potentiometers used in similar applications. The resolution can be set by the programming algorithm while the hardware itself remains the same. For antenna array beam optimisation a new effective synthesis method has been developed. This method is especially suitable for very large systems.

We were able to demonstrate the simulated performance of a model of a practical solid-state power amplifier. The experimental verification of the modelling work was beyond the scope of the project. However, the results, especially the phase response for the intermodulation products, were in line with what can be expected from this type of amplifier.

We also compared the indirect and direct predistortion learning architectures. In both the cases nonlinearities can be compensated if the amplitude range of the input signal is limited. The predistorter also tries to compensate the saturation area. It compensates harder this area if the amplitudes of the measured signal are higher. Consequently, we concluded that the result of the compensation depends also on the amplitude distribution, not only on the frequency. The direct learning architecture achieved a better performance in almost all cases. However, the results also showed that a direct learning model has a stability problem. In the indirect learning architecture ([R1]) the problem of commutability must be considered ([P2]). We pointed out that linearity is neither necessary nor sufficient requirement for the commutability. Furthermore, nonlinear systems are not in general commutative. Typical examples for commutability include (1) the systems are combined through a commutable operation, or (2) one of the systems is the inverse of the other system.

Finally, we presented a SystemC-based modeling approach for the dynamically reconfigurable hardware (DRHW). We used reconfigurable components mainly as co-processors in SoCs. The objective of the SystemC-based approach was to provide a mechanism that allows designers to easily test the effects of implementing some components in the dynamically reconfigurable hardware. The main contributions included an estimation approach for design space exploration and a model for reconfiguration overhead.

Publications

[P1] K. Masselos, N. S. Voros, Y. Qu, K. Tiensyrjä, M. Cupak, L. Rijnders, M. Pettissalo, "System level architecture exploration for reconfigurable systems on chip." In: Proc. 16th International Conference on Field Programmable Logic and Applications (FPL06), 2006.

[P2] A. Mämmelä, "Commutation in linear and nonlinear systems." Frequenz Journal of RF-Engineering and Telecommunications, Vol. 60, May/June 2006, pp. 92–94.

[P3] Y. Qu, J.-P. Soininen, K. Tiensyrjä, "SystemC-based design methodology for reconfigurable system-on-chip." In: Proc. IEEE 8th Euromicro Symposium of Digital System Design (DSD'05), 2005. Pp. 364–371.

[P4] Y. Qu, J.-P. Soininen, J. Nurmi, "A parallel configuration model for reducing the run-time reconfiguration overhead." In: Proc. IEEE Design, Automation, and Test in Europe (DATE'06), 2006. Pp. 965–970.

[P5] Y. Qu, J.-P. Soininen, J. Nurmi, "Using constraint programming to achieve optimal prefetch scheduling for dependent tasks on run-time reconfigurable devices." Accepted for publication in Proc. International Symposium on System-on-Chip (SoC'06), 2006.

[P6] A. Rantala, M. Sopanen, M. Åberg, "Implementation experiments of analog nonvolatile memory with a standard 0.35 µm CMOS." In: Proc. NORCHIP'04, 2004.

[P7] H. Salminen, P. Eskelinen, J. Holmberg, "An integrated differential single-chip VCO for S-band." In: Proc. European Frequency and Time Forum (EFTF'05), 2005.

[P8] L. Vaskelainen, "Constrained least-squares optimization in conformal array antenna synthesis." Accepted for publication in IEEE Transactions on Antennas and Propagation.

[P9] L. Vaskelainen, "Constrained least-squares optimisation in phase synthesis of conformal array antennas." Accepted for publication in Proc. First European Conference on Antennas and Propagation (EuCAP'06), 2006.

References

[R1] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials." IEEE Trans. Commun., Vol. 52, January 2004, pp. 159–165.

[R2] J. Kim, K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory." Electron. Lett., Vol. 37, November 2001, pp. 1417–1418.

[R3] D. Psaltis, A. Sideris, A. A. Yamamura, "A multilayered neural network controller." IEEE Control Systems Magazine, Vol. 8, April 1988, pp. 17–21.

[R4] "SICStus Prolog." Swedish Institute of Computer Science (SICS), [Online], Prolog development system, version 3.12.7, October 10, 2006. Available: HTTP: http://www.sics.se/sicstus.

13. New Filter Technologies – Development of Bulk Acoustic Wave Resonators and Filters

Jyrki Molarius, Arto Nurmela, Tuomas Pensala and Markku Ylilammi

13.1 Introduction

In a BAW (Bulk Acoustic Wave) resonator the acoustic wave is excited by applying a signal voltage across a piezoelectric layer (typically AlN or ZnO). This wave is confined to oscillate between the top air boundary and a bottom acoustic mirror structure. These interfaces reflect the longitudinal acoustic wave with a very high efficiency. The system is in mechanical resonance when the thickness of the vibrating structure contains an odd number of half waves. (Even-numbered harmonics leak through the acoustic mirror). The piezolayer converts the mechanical vibration back to electric signal and the resonator looks electrically like a series resonance circuit shunted by a capacitor equal to the capacitance of the dielectric piezolayer.

A filter is composed by connecting several resonators in either ladder or lattice topology. Also, more complicated acoustic structures can be used as filter. The quality of the BAW resonator can be characterized by two numbers. The electromechanical coupling coefficient k_{eff} determines the maximum bandwidth which can be attained in a filter, and the Q value determines the losses or power transmission in a filter. Usually the Q values are different for the series and parallel resonance frequencies. k_{eff} and Q are often combined in a figure of merit FoM = k_{eff}^2 Q. For good BAW filters to be used in RF applications in the 1–5 GHz range, k_{eff} should be over 0.23, Q > 1000 and FoM > 50.

The fabrication process of the BAW resonator consists of a sequence of thin film deposition and patterning processes on a substrate wafer. First the acoustic mirror is constructed by sputtering tungsten to a thickness equaling a quarter wavelength of the sound and patterning it by wet etch to prevent capacitive coupling between resonators. A good mirror requires two tungsten layers of high acoustic impedance (product of sound velocity and density) separated by chemical vapor-deposited silicon dioxide layers also in λ/4 thickness. SiO_2 has low acoustic impedance and as a dielectric it needs not to be patterned. The bottom metal electrode is patterned to form a contact area.

The piezoelectric layer is deposited in a magnetron sputter from target in reactive atmosphere (O_2 for ZnO, N_2 for AlN). The layer is patterned to allow contact to the bottom electrode. Then the whole structure is covered by a SiO_2 layer which is opened on top of the resonator and on the bottom electrode contact area. The last deposition step is to sputter aluminum with Mo adhesion and etch stop layer and to delineate it into top electrode and contact areas. Finally, the wafer is diced by sawing.

13.2 Deposition of the Piezolayer

Good quality piezoelectric layer is needed for Bulk Acoustic Wave (BAW) components. Earlier we have used sputter-deposited ZnO. Although a high coupling coefficient (k_{eff} = 0.28) was reached, results were not repeatable from run to run (i.e. wafer to wafer). Also the industry in the semiconductor field is only interested in AlN, because of its compatibility with the CMOS-process. Therefore, AlN was chosen as the piezoelectric layer for resonators and filters in this project.

The ZnO sputtering chamber was modified to accommodate AlN sputtering. Thermodynamically Al reacts faster with oxygen than with nitrogen forming unwanted aluminum oxide. Thus the growth of AlN is very sensitive to even low levels of oxygen contamination. This is well known in literature. Besides purity it is crucial to achieve right stoichiometry, stress, density, and correct crystal orientation (0002). To repel oxygen contamination heating elements were assembled in the deposition chamber and moisture and oxygen removing point-of-use filters were installed in the process gas lines (Ar, N_2).

With this set-up the AlN-resonator results were fairly good: k_{eff} 0.208 and Q values close to 1200 at series and over 600 at parallel resonance were reached. Repeatability was also much better than previously with ZnO. Although this performance could be satisfactory for some applications it is not enough for filters in radio transceivers. Therefore the effort was concentrated in improving the deposition process.

The distance between the target and the substrate was reduced from 115 mm to 65 mm (the minimum value possible in the chamber) to facilitate faster growth and to give higher energy to the particles hitting the growing surface. This should help in AlN growth itself and also reduce contamination. Substrate bias, which is used to tailor desired stress in the film, was changed from voltage control to power control for better stability (the power source handles this better). Oxygen levels in the AlN film went down to below 1% (0.7 at-% measured by RBS, Rutherford backscattering spectroscopy). An accepted value for the critical oxygen level which does not disturb piezoelectric quality is 0.5–1.0%. Despite this advance the coupling coefficient (k_{eff} = 0.20), the Q-values (Q = 650…1220) and the figure of merit (FoM = 45) did not markedly improve. Thickness uniformity over the wafer was now remarkably good (< 1.5%). These properties are not sufficient for constructing RF filters for mobile phone use, but may be able to serve some other applications.

To further accelerate the growth rate, a larger power supply (10 kW) was installed on the Al target. First trials with 3 kW power (9.55 W/cm^2) did not show a stable process. Lower power did not work too well either. If temperature was raised to enhance surface mobility, together with substrate bias for adjusting stress to slightly compressive side, substrate holder developed ground leaks. The location of shorts was pinpointed, however, and necessary modifications in the chamber were made.

The first results with 2 kW power using current regulation were not promising, (k_{eff} = 0.176, Q_s = 1200 and Q_p = 360) although compressive stress in the film was attained. Due to only a few runs in this last system configuration the full potential was not verified at the end of the project.

The main and practically the only reliable method to characterize the quality of a piezolayer is to fabricate a complete resonator and measure the electrical response. This is very tedious and slow. A much more efficient method was proposed which is based on the AFM (Atomic Force Microscope). This can measure the piezoelectric properties of a bare piezolayer on the substrate. This set up could be installed in the present AFM at VTT, but only the planning of that was completed in this project.

13.3 Development of the Resonator Structure and Fabrication Process

Besides the piezolayer also the structure of the resonator must be acoustically and electrically well designed. At high frequencies (1 GHz) the viscous losses in the metal electrodes appear to be a significant factor decreasing the Q value. An acoustic model (in Matlab) was developed to study the operation. Gold is an often used metal in BAW resonators but calculations show that it has all too high losses. Tungsten and molybdenum have low losses and they also produce high coupling coefficient. Ohmic losses are also important and their minimization was accomplished by combining molybdenum and aluminum in the electrodes.

A new processing step was introduced when the surface of the SiO_2 layer under the bottom electrode was polished by CMP (Chemical Mechanical Polishing) to be smoother than 0.3 nm RMS roughness. The high surface quality is expected to improve the piezolayer properties. The smooth surface must be maintained after depositing the bottom metal electrode. Several metals were tested. Electrically the best alternative Al grows rough and develops hillocks in heating. 120 nm of tungsten on top of Al prevents hillock formation. It was found that W, Mo, and Ni produce smooth surfaces on polished silicon dioxide. Electrically Mo is the best of these and it was chosen to be the bottom electrode.

In a filter component the metal layers in the acoustic mirror (SiO_2/W/SiO_2/W/SiO_2) must be patterned. This poses a difficult challenge to the CMP process. Usually polishing will strongly round the edges of the mesa structures of the mirror. However, applying a new FA (Fixed Abrasive) process resulted in smoothness better than 1 nm RMS without rounding the mesa edges. However, the thickness uniformity across the wafer still needs improvement.

BAW resonators may exhibit two possible types of acoustic behavior. The frequency – wave number dependence (dispersion) of the vibrating Lamb waves may be of type I or type II. In type I resonator the first TE resonance occurs at a higher frequency than the second TS resonance. In type II resonator the situation is vice versa. Resonators with ZnO piezolayer are always of type I, AlN resonators are intrinsically type II but the dispersion type may be switched by a suitable thin film stack design. Because resonators of dispersion type I can easily have a clean (spurious-free) response they are much more desirable than type II. The stack design was successful in producing type I dispersion.

In a filter component the piezolayer must be patterned in order to make electrical contacts to the bottom electrode. Because the piezolayer is thick (800 nm) compared to the metal electrode (200 nm) there is a severe problem how to make a continuous electrode line across the step in the piezolayer. Therefore attention was paid to develop a wet etching method for AlN which produces gently sloping edges. A formula based on phosphorus acid was found successfully etching slopes in 45 degree angles. There are still problems with point residue defects and poor selectivity against molybdenum and aluminum.

13.4 Filter Development

The development of a filter includes acoustic design of the BAW resonators and electrical design to connect them together to realize the desired transfer function. The filter can have ladder or lattice topology. The ladder filter may have balanced or unbalanced configuration, the lattice filter is always of balanced type. A new development was a design where a lattice filter have notches (transfer zeros) in the desired locations. This was accomplished by changing the resonator sizes.

Besides the above mentioned impedance element filters the RF filters can be realized by acoustic means. By stacking two piezolayers on top of each other it is possible to construct stacked crystal filters (SCF) and coupled resonator filters (CRF). These can give very wide pass bands and CRF is able to achieve balun-unbal and impedance level transformations in a very small size device. These were designed by the acoustic simulation model but no attempt to fabricate samples was made, as these structures are very difficult to process.

13.5 Spin-off Projects and Later Developments

Presently the packaging of BAW filters is the major part of the manufacturing costs of these devices. One major step to improve the situation would be to fabricate the packages for all filters (10000 /wafer) on the wafer before dicing it into chips. This can be accomplished by bonding a silicon wafer on top of the filter wafer to make the encapsulation. Development of this method was commenced in the project WASP (Wafer Scale Packaging) funded by TEKES and a German BAW filter manufacturer.

The application potential of thin film piezoelectric materials in acoustic structures is much wider than the simple BAW resonators. These possibilities are now investigated in the project MAKU (Microacoustics) funded by TEKES as a strategic research.

VTT was invited to participate in EU-project MOBILIS (Mixed SiP and SoC Integration of Power BAW Filters for Digital Wireless Transmissions) where the task of VTT is to develop BAW resonators for high power applications.

The know-how of VTT in the field of BAW devices has been utilized in projects ordered by German (project IBAW), Japanese (project MEWBAW) and Netherlandian (project QoT) semiconductor manufacturers.

Publications

[P1] J. Molarius, J. Kaitila, T. Pensala, M. Ylilammi, "Piezoelectric ZnO films by r.f. sputtering." Journal of materials science: Materials in electronics, Vol. 14, (2003), No: 5–7, pp. 431–435.

[P2] T. Pensala, T. Makkonen, J. Vartiainen, J. Knuuttila, J. Kaitila, O. Holmgren, M. M. Salomaa, "Laser interferometric measurement of Lamb wave dispersion and extraction of material parameters in FBARs." Proceedings of the 2002 IEEE International ultrasonics symposium, München, DE, 8–11 Oct 2002. IEEE Ultrasonics, Ferroelectrics and Frequency Control Society (UFFC-S), 2002.

[P3] J. Kaitila, M. Ylilammi, J. Ellä, R. Aigner "Spurious Free Bulk Acoustic Wave Resonators." Proceedings of the 2003 IEEE International ultrasonics symposium, Honolulu, US, 5–8 Oct 2003. IEEE Ultrasonics, Ferroelectrics and Frequency Control Society (UFFC-S), 2003.

[P4] T. Makkonen, T. Pensala, J. Vartiainen, J. V. Knuuttila J. Kaitila, M. M. Salomaa, "Estimating Materials Parameters in Thin-Film BAW Resonators Using Measured Dispersion Curves." IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 51, No. 1, 2004.

14. 4G Prestudy

Markku Åberg and Silja Vatunen

The goal of the 4G Prestudy project was to make a research plan for a fourth generation (4G) mobile terminal radio development research project. The target 4G research project will be a major strategic (theme) project within VTT and its plan will help to focus the actual thematic programme by providing a frame for the VTT's 4G research efforts.

The tasks of the project were:

Task 1. Survey of the existing scenarios and roadmap
Task 2. Survey of existing/planned major projects in the field
Task 3. VTT 4G roadmap
Task 4. Detailed 4G research project plan.

Tasks 1, 2 and 3 were covered by an extensive literature survey and discussions with key customers.

As a result a summary ”Technical Report; Prestudy for 4G Radio research project” was written. The main themes suggested are (not in any order of preference):

1. Study of multi radio architectures
2. Functional antennas (as an integral solution with adaptive matching, switches and possibly low noise amplifiers)
3. Discrete time radio architectures
4. HW-SW co-design in adaptive multi radio systems.

Availability of advanced MEMS and sub-100 nm CMOS processes is a key factor in all themes.

A project proposal about the selected themes mentioned above was prepared for the VTT theme program “Digital World”. with the title “Flexible Wireless Platforms for 4G Systems” (WIPLA). The proposal was evaluated by external evaluators as one of the best submitted for the theme, and got the applied financing, starting on August 1, 2006.

15. Enabling Technologies for Smart Human Environments (ETSHE)

Jarmo Keski-Säntti, Tapani Rantakokko, Ari Virtanen, Johan Plomp, Veijo Lappalainen, Kari Haataja, Vesa Kyllönen, Lauri Pohjanheimo, Jani Mäntyjärvi, Heikki Keränen, Jouni Kaartinen, Veli Möttönen and Kalevi Piira

15.1 Project Goals

The encapsulated goal of the project was the realisation of the vision of a smart human environment, which required and still requires advancements in communication technologies, interaction technologies and ambient intelligence. Because of the participation of four VTT groups one essential target was combining different fields of know-how as integrated result. Those groups participating for this project were from VTT Electronics, VTT Industrial systems and VTT Building technology. One of the earliest goals was scenario work, for describing ideas of future environment, which was performed together and as co-operation with NISHE project.

The goal of the project was to identify, develop and demonstrate technologies that will become essential in building the smart human environments of the future. A smart human environment can be seen as the combination of various sensing and communicating objects that make up a smart office, home, car or other physical entity, and support human users within the environment in their various tasks.

The building technology research aimed to provide viewpoint of building services and maintenance and provide the building service application module. Also the scenarios concerning building services in home and office environments and an integrated smart human environment service and system model focused to home and office environments.

Ambient intelligence and positioning technologies alone are difficult to demonstrate separately. Accordingly, one of the goals was to demonstrate location aware, intelligent services in many kinds of environments e.g. home, office, factory or outdoors. As goal for that was developed positioning utilizing which included both indoor and outdoor positioning methods. Seamless integration towards the system, which can be called as seamless navigation was targeted. Navigation and guidance methods for inaccurate maps and positioning was planned to be realized. These techniques were possible only with equipments which can utilize many communication protocols.

However, smart human environment does not need to be restricted only to physical places and combinations of physical objects – it can follow the user as he/she is on the move, i.e. be mobile, and utilize services that are not bound to physical objects, i.e. be (at least partially) virtual. The project stresses the human side of the smart environment: The environment is not supposed to be smart "as-is" and be able to act on behalf of humans. Instead, the technology is supposed to let humans act smarter by helping wherever possible. Therefore, the enabling technologies for smart human environments may consist of for example sensing hardware, algorithms, communication software, data storage, user interfaces etc.

15.2 Project Implementation

At the beginning of the research the borderlines were sketched together in workshops in order to get understanding on common goals and to make the scenario document. The scenario was planned and realized iteratively in co-operation between the participating researchers and with an another VTT project called NISHE (Natural Interactions in Smart Human Environment).

In practice the work in this project was divided into several work packages. WP1 and WP7 were joint work packages where all partners participated. WP1 was planned to co-operate also with the NISHE project for the definition of scenarios and implementation of pilot services, but because the NISHE project ended after 2002 the implementation and pilot services were not performed together. Work packages WP2 to WP6 are technology research work packages mostly conducted by the partner specialised in that particular matter. The contents of the packages were flexible and the research topics were focused during the project work after evaluating preliminary results. Those work packages were:

- WP1 Scenarios and pilot services
- WP2 Ambient intelligence platform
- WP3 Ambient intelligence methodologies
- WP4 Indoor positioning techniques & navigation
- WP5 Indoor and outdoor positioning integration
- WP6 Service models.

The project was performed during the years 2002 and 2003.

15.2.1 Ambient Intelligence Platform

Acknowledging the rise of the number of mobile terminals and information exchange within the various virtual communities of the Internet, the focus was placed on a few promising, emerging technologies that support more versatile communication between people via the means of personal mobile devices.

This approach was chosen because the communication between people is an essential part of practically any human environment, and thus enabling convenient and efficient communication is one of the basic features expected also from a smart human environment. It is useful to note that communication may occur in real-time such as conversations over phone, but also “offline” via guiding signs, bulletin boards etc. Novel technologies can assist humans in all kinds of communication.

The theory of the research was based on context awareness, which claims that devices will be more useful to humans when they collect information about the circumstances under which they operate and react accordingly.

For example, a device that is able to utilize smart sensor technologies for collecting low-level information of its own whereabouts, orientation and movement, can process the data to human-readable format and perhaps visualize it with a map and a few graphical icons. These small pieces of information may in turn be extremely useful when shown to another person who is about to communicate with the user of the sensor-enabled device. From a bunch of simple clues a human can make smart assumptions, like whether it is appropriate and reasonable to place a voice call or send an email instead. Thus, by providing humans with useful information that is not otherwise easily available, the communication between people becomes more convenient and smarter, and thus facilitates smart human environments.

On the other hand, a human environment that utilizes technology to make it smarter lies at the boundary of a human and a computer. Therefore some kind of a user interface is a necessity. As mentioned earlier, smart human environment may provide services independent of current location, which makes personal mobile terminals suitable for presenting the user interface and handling user interaction. Mobile terminals can also be equipped with various sensing mechanisms, which provide a lot of information as the measurable variables of the environment change whenever the user moves (assuming he or she is carrying the terminal along). Thus focusing on mobile terminals is well grounded.

Since the goals of the project included developing and demonstrating novel technologies for smart human environments, a prototype system consisting of several mobile terminals enhanced with sensors and special software were to be created, demonstrated and published.

15.2.2 Ambient Intelligence Methodologies

The methodologies for decision making in the smart environment were studied based on literature and with co-operation of other work packages and part of them in piloting. The decision making methodology concentrated especially on novel, but simple and fast calculation techniques which can be utilized also in environment that have small computing capacity. The communication language and configuration problems were taken into account.

The importance of location knowledge was recognized as the most useful context reasoning type at the early stage of the evolution of the ubiquitous environment. Location data can be utilised for location based services, maintaining tasks, safety reasoning and for many other purposes in ubiquitous environments. Therefore location reasoning is chosen for the pilot reasoning software module. For the pilot purpose, the WLAN positioning was utilised and indoor route planning pilot program was developed. Also utilisation of physical pointing was studied at this work package.

15.2.3 Indoor and Outdoor Positioning Techniques, Navigation and Integration

This work packages were performed in co-operation with a research project NOPPA.

The utilisation of indoor and outdoor positioning methods was evaluated. Navigation and guidance methods for inaccurate maps and positioning were developed. A brief study of indoor navigation for the purpose of implementing a suitable demo was made. A generic algorithm with inputs of positioning and heading information from different outdoor and indoor positioning systems can produce a single positioning and heading information for navigation.

The implementation was part of a larger VTT's Information server concept which aimed to utilise commercial services and devices for improving public transport accessibility with creating access to passenger information with a personal mobile device rather than building physical infrastructure.

15.2.4 Service Models

The common pilot was defined in co-operation with researchers; the definition included the concept and architecture of overall demonstration as well as the definition of pilot modules. The pilot modules were implemented separately by related teams.

The building services scenarios of home and office environments were thematically created in co-operative workshops and written individually according to the agreed distribution of work.

The development of the services model of smart human home and office environments was carried out by teamwork organising several workshops and developing the ideas forward in between times.

15.3 Description of Results

The scenario document of smart human environment was created in co-operation with NISHE-project. It was written in Finnish and called *Intelligent Consumerism in Shopping Mall,* including five different sub-scenarios. The main scenario included a fictional journey from home to a shopping mall during an imagined weekend. At the scenario there are about 50 ideas, which were noted to be interesting in future, but were not yet practical. Seamless communication techniques, guiding, notebooks, service, visual, ubiquity, and communal issues were among the ideas. Also five sub-scenarios were included on positioning, sound recognition, visual tags and video.

In the sub-scenario *Behind the Scenes* there is case when a service man comes to repair a damaged building service system equipment. The scenario includes the alarm detection and notification to the service centre, guiding of the service man to the building, to the right door of the parking hall and to the parking place as well as to the equipment inside the building After coming to the equipment the internet enabled mobile user interface assist the service man showing all relevant product, operation and maintenance information of the equipment with an easy user interface. The systems takes care of life-safety issues like connecting electricity supply off before the service work. When the work is done the system logs the required information for service history and cost control.

15.3.1 An Adaptive Map-based Prototype for PDA

The purpose of the prototype was to allow experimenting with technology enhanced communication. As the project was technology oriented, the chosen approach was to add a versatile set of sensors to a commercial PDA terminal device, to create software drivers and algorithms for acquiring sensor data and turning it into formats that are useful for humans, and to identify and to develop applications that could facilitate communication with sensor based information.

The sensor part was based on a so-called SoapBox device, developed at VTT in earlier research projects. Each PDA was equipped with one SoapBox that was attached into the back of the terminal and connected to it with a short serial cable (Figure 1). The SoapBox provided various sensors, including 3D accelerometer, 2D magnetometer, light intensity, and proximity. In addition, the PDA had a microphone for recording the soundscape of usage environment and wireless network connection for communicating with other terminals and computers.

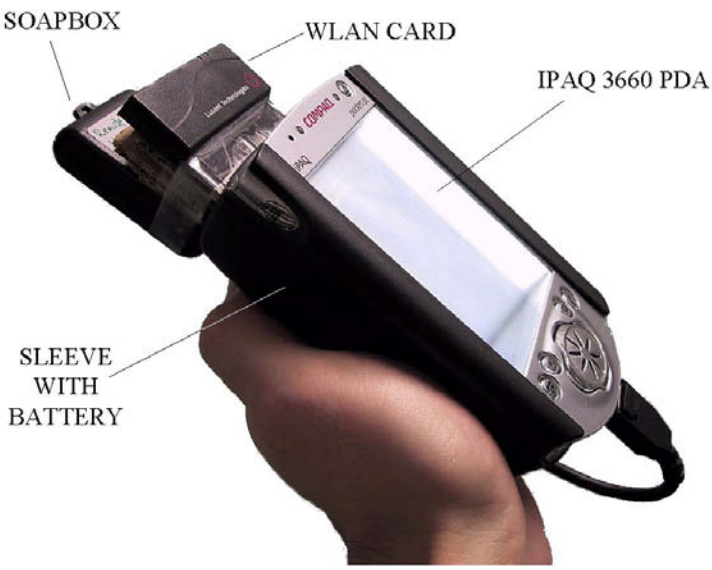


Figure 1. The PDA terminal enhanced with SoapBox sensors.

By the means of the developed algorithms, the sensor information was turned into physical user activity (WIFI based positioning, accelerometer based movement, accelerometer + magnetometer based orientation), a set of environment variables (microphone based noise level and light intensity sensor based lightness level), as well as device activity (current application or task, active or idle). A set of simple graphical icons were designed for visualizing context information with ease on small screens (Figure 2).

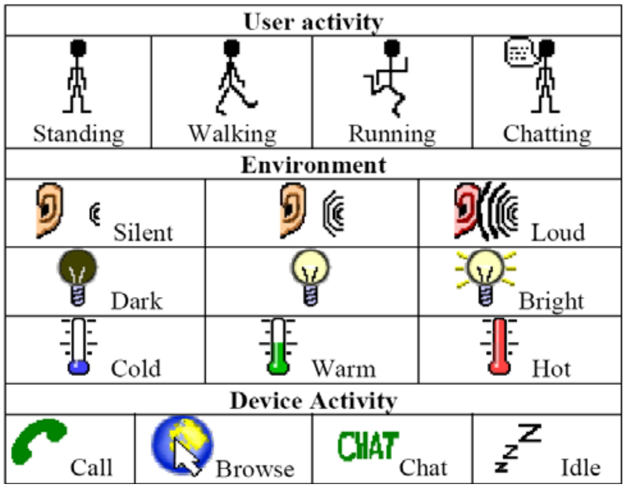


Figure 2. Graphical icons were designed for visualizing context variables.

A shared map application for a group of users was developed for experimenting and demonstrating communication enhanced with technology. Each user of the group runs personal copy of the map-application in his/her terminal. The content shown on the map is synchronized between users automatically in the background.

The map-application renders an XML format map to vector graphics, and supports panning, zooming and rotating features for map manipulation. Location of each user of the group is visualized on the map using an icon (avatar). By selecting a user icon, one can see the current context of that user including physical activity, environment and device activity variables.

One potential use case for such application could be a group of urban youngsters who create multimedia content using terminal’s sound, photo and video recording capabilities, and then share the media clips with their friends (other users of the group). The benefit of the application would be that instead of actively sharing content via sending messages to each individual via emails or MMS messages, the user who is creating the content does not need to perform any steps at all for sharing the media clips. They will become available for others in the group via automatic synchronization of map objects: the content simply shows up on the map as clickable objects in the location where they were initially created. See Figure 3 for an example based on an ETSHE prototype application. This prototype was also tried with navigation help and positioning utilisation.

Figure 3. Screenshots describing a scenario in which Joe has recorded a short video clip of a stand-up comedian and placed the document on the map. Arrows between screens describe navigation achieved by clicking at the starting point of the arrow.

15.3.2 NOPPA Architecture

At the utilisation of indoor and outdoor positioning methods the ETSHE project was performed at the same time as the NOPPA project. Actually, it can be said that the ideas and techniques developed in ETSHE are practiced in NOPPA, so here we present the NOPPA architecture as a result.

VTT's Information server concept aims to utilise commercial services and devices for improving public transport accessibility with creating access to passenger information with a personal mobile device rather than building physical infrastructure.

Research and development of the concept was done in a research project NOPPA (Navigation and guidance for the visually impaired), which started in June 2002 and ended in 2004. The project piloted a personal navigation system producing a unbroken trip chain for the visually impaired.

The most important building blocks of the system are the mobile Internet, mobile phones and personal navigation systems and services. For an unbroken trip chain for the visually impaired there are requirements for continuous, general use positioning techniques, continuous (Internet) access to passenger information and availability of map data accurate enough for the users' needs. For example there are not generally available maps including pavement information or information about large public premises. A door-to-door guidance would require map data including entrances, and continuous guidance would even require indoor maps and indoor positioning, which are generally unavailable. The design goals of the NOPPA guidance system were:

- Easy and fast to use
- Flexibility
- Affordable to the user
- Access to public transportation and passenger information systems
- Applicable both indoors and outdoors
- Integration of products and services for personal navigation
- Modular, easy to update, easy to add functions
- Speech user interface.

To fulfil these requirements an architecture presented in Figure 4 was created. Because of low processing capacity of the mobile terminal and a low bandwidth wireless data connection, most of the work is done in the Information server. Data flow between mobile client and Information server is minimized, which keeps the communication costs low and shortens response time.

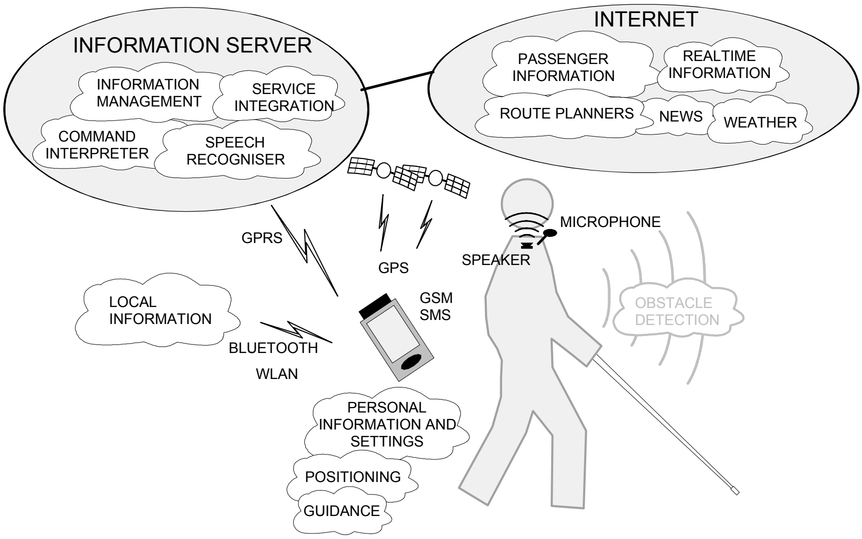


Figure 4. Architecture of the NOPPA guidance system.

15.3.3 Building Service Pilot

The idea of the building service related pilot was the following: There is a fault in an air handling unit and it alarms to the remote service provider. The service man does not know the building. First he has to find the right building, right room and the device by using his mobile device. After that he gets the service history and instructions from the back systems. After the preparation operations is done he electrically signs the job to the back system. Its implementation consisted of the following features and parts: building faults, alarms and sensor (temperature, CO2, …) information emulator with internet gateway (Lonwork + IP gateway based demonstration bag), the life cycle database of air handling unit (example of the service history data from web based background system), interface to WLAN based localization information inside the building (other partners of the project – not implemented into the building service related pilot) and mobile user interface (web based interactive CAD –pictures).

The building services scenarios of office and home environments (in Finnish) includes the following scenarios. The scenarios of the home environment were the following: entering the gate and garage, entering from outdoor, leaving home, entrance to the room, exit from the room, being in the room, going to sleep, sleeping, waking up, cooking, going to the sauna, laundring, other homework and visit. The scenarios of the office environment were following: entrance to the office, exit from the office, entrance into the office room, exit from the office room and working in the office room.

The results of the development of the services model of smart human office and home environments consists of the following working documents: the description of the technical systems solutions related to home and office scenarios (in Finnish) and two PowerPoint presentations. The theme of the first one was the service model – background analysis and integrated service modelling approaches. The themes of the other one were Facilities Management (FM) and Smart Human Environment, FM Systems development, Developments in Building Modelling, What Next in FM, Ubicomp / Smart Environment, Bridging Physical and Virtual Worlds. The work was focused to the preliminary analysis of the service modelling approaches, which proved to be very demanding and extensive.

15.4 Conclusions

The scenarios were developed according to the project plan. They were good starting points in composing the pilot ideas and modelling approaches. The main parts of the pilot was implemented according to the plan but the integration part with the other partners was not done because the integration would have needed more time and the real development and testing environment of the WLAN based localization system (inside the building) was in a different geographical site. One additional problem with piloting was the unexpected end of the NISHE project.

The research on technologies for smart human environments was focused on enhancing communication between humans via the means of mobile terminals equipped with various sensors. Prototype devices were created for experimenting and demonstration purposes. The prototypes consisted of PDA terminals equipped with VTT's own SoapBox sensor add-ons. Drivers, algorithms and graphical icons were developed for collecting, abstracting and visualizing context information on small screens. A map application was developed for enabling context aware communication and sharing of multimedia content between a group of users.

The results of the project were published in many conferences and used in academic theses, including one diploma thesis and a PhD thesis. The diploma thesis got SIGCHI Finland's award for the best academic thesis in the field of human-computer interaction in 2003. The developed prototype was frequently demonstrated at VTT and later used in several user studies in a three year ADAMOS project, funded by the Academy of Finland.

Since the time of the development of ETSHE prototype and concepts, sensor based context information has been developed further in the research community and applied in mobile terminals by various manufacturers. For example, built-in 3D accelerometers have recently appeared in Nokia's 5500 sport phone model. The phone uses the sensor as a step counter, while the ETSHE prototype demonstrated the recognition of standing/walking/running modes. Many phone models use light intensity sensor for adjusting display brightness and thus save battery power.

Context information is now also emerging for communication between people. So-called context-aware or presence enabled phonebooks have been demonstrated in mobile phones and are already being used in instant messaging applications. The phone books, for example, visualize the current context of each person in the phonebook, including such as current location and device activity. Graphical icons are used for representing the information. This is very similar to the ETSHE prototype.

Maps are also becoming more commonplace in mobile terminals along with navigation applications supported by GPS positioning. Sharing multimedia content such as photographs by linking them on maps is now possible in popular Internet services such as Google Sightseeing and Flickr.

The most important value of the ETSHE prototype was that it combined such a multitude of novel features into a one, functional prototype in a time when these technologies were still very young. The multi-functional prototype enabled creating applications that efficiently demonstrated the benefits and possibilities of enhancing mobile terminals with sensors, and thus paved the way for technology-assisted communication in the smart human environments of the future.

The NOPPA prototype shows that a guidance system for the visually impaired person is possible to be build without laborious and expensive changes in the current infrastructure. This is possible by utilising commonplace Internet services for passenger information and personal navigation. It seems applicable to at least buses, trains and trams in the same format. In air traffic the flight and gate status can be followed.

Passenger information systems should be equipped with a standard Internet interface (XML). Through this interface it would be possible to deliver static information (timetables), real time information and disturbance information in such way that it is accessible to the special user groups. These interfaces are needed for example when developing mobile applications for all passengers. By utilising real time information systems it is possible to guide visually impaired persons to right vehicles without costly investments to the beacon systems

The implemented pilot module at the building research areas proved the validity of the concept and it has been useful starting point for many related projects. Instead of the final service model of smart home and office environments, only the different modelling approaches and methodologies were analysed and developed.

Publications

[P1] H. Keränen, T. Rantakokko, J. Mäntyjärvi, "Sharing and presenting multimedia and context information within online communities using mobile terminals". Proceedings of International Conference on Multimedia and Expo, ICME '03. Vol. 2. Baltimore, MD, USA, 6–9 July, 2003. IEEE, 2003. Pp. 641–644. http://ieeexplore.ieee.org/servlet/opac?punumber=8655.

[P2] S. Koskinen, A. Virtanen, "Navigation System for the Visually Impaired Based on an Information Server Concept." In: Proceedings of the MobileVenue'04. 27–28.5.2004 Athens, Greece. 6 p.

[P3] S. Koskinen, A. Virtanen, “Public transport real time information in Personal navigation systems for special user groups.” In: Proceedings of 11th World Congress on ITS. 18–24 October 2004, Nagoya, Aichi, Japan.

[P4] J. Mäntyjärvi, T. Seppänen, “Adapting applications in mobile terminals using fuzzy context information.” Interacting with Computers Journal, Vol. 15(3), Elsevier, 2003.

[P5] J. Himberg, J. Häkkilä, J. Mäntyjärvi, P. Kangas, “On-line personalization of a touch screen based keyboard.” In: Intl. ACM Conference on Intelligent User Interfaces, Miami, USA, Jan 2003.

[P6] J. Mäntyjärvi, J. Himberg, P. Huuskonen, “Collaborative Context Recognition for Handheld Devices.” In: IEEE Intl. Conference on Pervasive Computing and Communications, Dallas, USA, March 2003.

[P7] J. Mäntyjärvi, U. Tuomela, I. Känsälä, J. Häkkilä, “Context Studio – Tool for Personalizing Context – Aware Applications in Mobile Terminals.” Computer-Human Interaction Special Interest Group (CHISIG) of the Ergonomics Society of Australia and New Zealand, Oct 2003.

[P8] J. Himberg, J. A. Flanagan, J. Mäntyjärvi, “Towards Context Awareness Using Symbol Clustering Map.” Workshop on Self-Organising Maps, Kitakyushu, Japan, Sept 2003.

[P9] S. Kallio, J. Kela, J. Mäntyjärvi, “Online gesture recognition system for HCI of mobile terminals. IEEE International Conference on Systems.” Man and Cybernetics, Oct 2003.

[P10] T. Rantakokko, “User interface adaptation based on context.” Oulu, University of Oulu, Department of electrical engineering, Masters Thesis, 2003. 84 p.

[P11] T. Rantakokko, J. Plomp, “An adaptive map-based interface for situated services.” Smart objects conference, sOc 2003. Grenoble, FR, 15–17 May 2003. Pp. 218–221.

[P12] A. Virtanen, S. Koskinen, "Towards Seamless Navigation." In: Proceedings of the MobileVenue'04. 27.–28.5.2004 Athens, Greece. 6 p.

[P13] A. Virtanen, S. Koskinen, "Information Server Concept for Special User Groups." In: Proceedings of 11th World Congress on ITS. 18–24 October 2004, Nagoya, Aichi, Japan.

[P14] A. Virtanen, S. Koskinen, "Noppa, näkövammaisten opastusjärjestelmän pilottiprojekti." Loppuraportti. VTT Tuotteet ja tuotanto. Tammikuu 2005. 126 s.

[P15] A. Virtanen, "NOPPA, Navigation and Guidance System for the Visually Impaired. Societal Applications and Repercussions of Language Technologies." Summer Course. The Universitat Internacional Menéndez Pelayo – Centre Ernest Lluch, Barcelona, Spain, 11.–13.7.2005.

[P16] A. Virtanen, E. Chandler, J. Kavanagh John, "Accessible Bus Timetables and Real Time Information." Techshare 2005. 17–18 November, 2005, Birmingham, UK.

[P17] A. Virtanen, "Mobile devices and future trends." Techshare 2005. 17–18 Nov 2005, Birmingham, UK.

[P18] A. Virtanen, "Mobile assistant for disabled – and any – public transport customer." InformNorden 2006. 14.–15.5.2006. Stockholm, Sweden.

[P19] A. Virtanen, "Henkilökohtainen navigointi näkövammaisen apuna." Positio 2/2002. Maanmittauslaitos, Helsinki 2002.

[P20] A. Virtanen, "Noppa, näkövammaisten opastusjärjestelmä". Esitelmä. Ikääntyvä Suomi – Miten rakennamme kaikille soveltuvan tietoliikenne-ratkaisun? – Seminaari, Stakes, 22.8.2002.

16. Context Sensing with Wearable Sensors and Data Fusion (Palantir)

Juha Pärkkä, Luc Cluitmans, Miikka Ermes,
Mark van Gils, Ilkka Korhonen, Panu Korpipää, Jani Mäntyjärvi,
Johannes Peltola, Jouko Vilmi, Pasi Välkkynen and Arto Ylisaukko-oja

16.1 Introduction

The objective of the Palantir project was to study how and which meaningful contexts it is possible to recognize from data collected with wearable sensors. With context we mean any information that describes the surroundings or situation, e.g., of a portable device. When developing context sensitive systems, the aim is to give a device senses. With help of these senses the device will be capable of collecting information about its surroundings and situation. Different devices can then use the context information in different ways, e.g., for automatically keeping a digital diary, for automatically adapting the user interface or user profile, for automatically recommending a service or information for the user, etc.

The approach of this project to context sensitivity was data-oriented and empirical. A wearable data collection system was developed to allow realistic context data collection with several wearable sensors, from several different realistic contexts by several volunteers. The data collection resulted in a large context data library. New methods for context detection were developed by analysing and combining the different collected signals. As a result of this project, we identified which contexts can be realistically recognized and which sensors and data processing methods are the most useful ones for this purpose.

This study continued VTT's research on ubiquitous computing and smart human environments (SHE). The participating research groups had previous experience in context sensing, signal processing & analysis and wireless sensor networks (SoapBox).

The Palantir project ran from June 2003 till December 2004 and it received funding from VTT Future Communications Technologies technology theme, Tekes Fenix programme and from the participating companies: Nokia, Clothing+ and Suunto. The research partners were VTT Information Technology and VTT Electronics.

16.2 Target Contexts

The starting point was to focus on a person doing outdoor activities and measure contexts with wearable sensors. These assumptions lead us to focus on sensors measuring the **human** and the **environment**. We also chose to do **long-term measurements in out-of-lab, real-life conditions**. The contexts we are interested in are:

- Activity (sitting, standing, …),
- Location (indoors, outdoors, vehicle, …),
- "Gaps" in life" (waiting, traveling),
- Social contexts (alone, in a crowd),
- Eating (eating, drinking),
- Health (coughing, sneezing, …).

With **Activity** contexts we try to identify, how accurately it is possible to measure the level of activity or even some specific activities like different sports. The goal is to have an application that would give statistical information on the activities the person has done during a day or week or a longer period. The application could tell: "Today you have been sitting 32%, walking 21%, standing 24%, lying 13%, exercising 10%". If specific actions or sports can be identified reliably, they could be represented with time: "Today you have been playing football for 48 minutes".

By measuring **Location** contexts, our aim is to detect how much the person has been indoors, outdoors or in a vehicle. In applications, this information can be used in many ways: for statistical information, for immediate UI adaptation etc. The statistical information would be interesting to anyone: "How much of your time did you spend traveling this week?". "How much of your time did you spend outdoors this week?" etc.

With **“Gaps in life”**-contexts we are interested in identifying the moments the person has nothing to do and he would be willing to use a mobile gadget for news, entertainment, etc. In our scenario for data collection, waiting for bus and traveling by bus are such moments.

With **Social contexts** we are interested in identifying, if the person is alone or in a crowd. This information could be interesting as statistical information too, but more useful the information might be, e.g., for updating the UI in real time (like ringtone of a mobile phone).

With **Eating pattern** context we are interested in identifying the eatings and drinkings of a day. This is also statistically interesting information: “How many times per day do you normally eat/drink?” In real-time applications this is also important information, e.g., many people don’t want to be disturbed when eating.

With **Health** context we are interested in detecting the coughing and sneezing. These could be interesting as statistical information in some special circumstances. However, for real-time use detecting coughing would be interesting e.g. when detecting and warning about altitude sickness.

16.3 Data Acquisition System

The data acquisition system was developed during August – October 2003. It measures all together 18 different quantities from the user and his environment: 1) altitude, 2) audio, 3) body position, 4) chest accelerations (3D), 5) chest compass (3D), 6) ECG, 7) environmental humidity, 8) environmental light intensity, 9) environmental temperature, 10) heart rate, 11) location, 12) pulse plethysmographic waveform (PPG), 13) respiratory effort, 14) blood oxygen saturation (SaO2), 15) skin resistance, 16) skin temperature, 17) wrist accelerations (3D), 18) wrist compass (2D).

Some of these quantities are measured using many sensors. E.g., acceleration is measured in 3 dimensions, thus resulting in 3 channels of data, heart rate is obtained from finger and forehead oximeters, from Suunto X6HR heart rate belt and it can be derived from ECG. Thus the measurement resulted in 35 channels of data. During measurement sessions, the data were stored into four devices:

1. solid-state-memory based recorder (Embla A10, Medcare, Reykjavik, Iceland),
2. rugged, small-size PC (Databrick III, Datalux, Winchester, Virginia, USA),
3. wrist-top computer for measurement of heart rate and altitude (Suunto X6HR, Suunto Oy, Vantaa, Finland), and
4. GPS navigator for recording location and altitude (Garmin eTrex Venture, Garmin Ltd, Olathe, Kansas, USA).

The user carried the sensors and other equipment on him and in a normal rucksack. With the equipment the rucksack weighed about 5 kg. After the measurement session, the data was uploaded to a PC for off-line data processing and analysis. The data acquisition system is described in more details in Figure 1 and Table 1.

SensorBox is a prototype equipped with several sensors. It measures both environmental and physiological signals from volunteer's chest as it is attached to the rucksack strap. **SoapBox** (Sensing, Operating and Activating Peripheral Box) is a small, low-power-consumption device that includes a microcontroller, serial port, bidirectional short range radio communication and sensors. In Palantir, it measured 3D acceleration and 2D magnetometer signals on wrist.

Figure 1. Palantir Data Acquisition System: Left: equipment on the floor, Right: equipment in the rucksack. Left, topmost row: Rucksack. Second row: PC battery, PC, Embla recorder (white) and battery. Third row: Oximeter (gray), SoapBox (black), microphone and amplifier (black), SensorBox. Fourth row: Respiratory effort strap (blue), body position (blue, around strap on the left), ECG electrodes (white-blue), skin temperature (silver, on top of blue strap), skin resistance (gray, around blue strap), Suunto wrist-top computer and its heart rate strap below, second oximeter (black), GPS navigator.

Table 1. Palantir Sensory Equipment.

Signal	Sensor	Measurement site	Storage	Sampling rate [Hz]	Info	Device
Altitude	Air Pressure	Wrist	Suunto X6HR	0,5	Ascent, descent	Suunto X6HR
Audio	Microphone	Chest, on rucksack strap	Databrick PC	22000, mono, 16 bit	Audio	AKG C417
Body Position	Metal ball moves between resistors	Chest, around respiratory effort belt	Embla	200	Lying on back/ stomach/ right side/ left side	Pro-Tech Position
Chest Accelerations	3D acceleration	Chest, SensorBox	Embla	200	Movements, Position	SensorBox, 2 x Analog Devices ADXL202
Chest Compass	3D compass	Chest, SensorBox	Embla	200	Body Direction	SensorBox, Honeywell HMC1023
EKG	Voltage between EKG electrodes	Left armpit on upper breastbone	Embla	200	Heart activity	Blue Sensor VL
Environmental Humidity	Humidity	Chest, SensorBox	Embla	200	Humidity	SensorBox, Honeywell HIH-3605-B
Environmental Light Intensity	Light sensor with two output dynamics	Chest, SensorBox	Embla	200	Light intensity	SensorBox, Siemens SFH203P
Environmental Temperature	Temperature sensor	Chest, SensorBox	Embla	200	Temperature	SensorBox, Analog Devices TMP36
Event Button	Switch	Chest	Embla	-	Annotation for Coughing and sneezing	Embla XN Oximeter with Event Button
Heart Rate	IR light absorption	Finger	Embla	1	Heart Rate	Embla XN Oximeter
Heart Rate	IR light reflectance	Forehead	Databrick PC	3	Heart Rate	Nonin XPOD oximeter
Heart Rate	Voltage between belt electrodes	Chest, HR belt	Suunto X6HR	0,5	Heart Rate	Suunto X6HR
Location	GPS satellite receiver	Shoulder, on rucksack strap	GPS receiver	Based on location	Location, speed	Garmin eTrex Venture
PPG (pulse plethysmographic waveform)	IR light reflectance	Forehead	Databrick PC	75	Perfusion, blood volume changes	Nonin XPOD Oximeter
Respiratory Effort	Piezo sensor	Chest, elastic respiratory effort belt	Embla	200	Respiratory movements	Pro-Tech Respiratory Effort
SaO2	IR light absorption	Finger	Embla	1	Blood oxygen saturation	Embla XN Oximeter
SaO2	IR light reflectance	Forehead	Databrick PC	3	Blood oxygen saturation	Nonin XPOD Oximeter
Skin Resistance	Resistance between two parallel metal leads	Chest, around respiratory effort belt	Embla	200	Sweating	Custom VTT sensor, DC
Skin Temperature	Resistive temperature sensor	Back, below neck	Embla	200	Exercise, Health	YSI 409B
Wrist Accelerations	3D acceleration	Wrist	Databrick PC		Hand movements, position	VTT SoapBox
Wrist Compass	2D compass	Wrist	Databrick PC		Hand direction	VTT SoapBox

Start and end markers were manually added to all data to allow synchronization of devices. The manual markers achieved an accuracy of +/-0,5 sec. This was considered adequate in such long recordings (2 hours and 6 hours). Two software applications were written for acquiring data from oximeter, microphone and SoapBox on PC.

16.4 Annotation System

A PDA based annotation application was developed for the annotator to mark the true context. The application was written using C#.NET and it runs on platforms that have .NET framework installed (e.g., PocketPC and Windows with .NET installed). The application visualizes the currently selected (active) contexts and provides a UI for changing the selected context. The UI can be easily edited as the layout is defined in a separate xml-file.

The application organizes contexts into upper level *context types* (activity, location, …) and lower level *context values* (sit, lie, indoors, …). Only one of the context values can be active at a time. As the volunteer's context changes, e.g., he sits down, the annotator taps on the new context value ("sit") with the PDA pen. The application stores the new state with a timestamp on PDA memory. The xml-based log-file of annotations can be downloaded to PC and analyzed there.

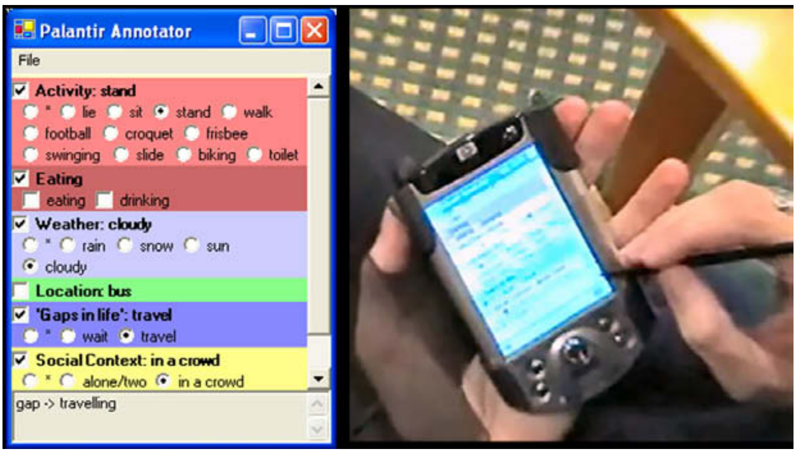


Figure 2. Palantir Annotation Application for storing the "true" context.

16.5 Data Collection

In data collections the idea was to collect realistic context data in out-of-lab settings. The aim in 2003 was to perform long measurement sessions of 2–3 hours with 10–20 volunteers. In 2004 data collection the aim was to measure circa 6 hours with 10–15 volunteers. Long recordings of context data make it possible to see how the context data behaves in real-life conditions and to base the algorithm development on realistic data.

16 volunteers took part in 2003 data collection and 12 volunteers in 2004. The reason for collecting context data with so many persons lies in the fact that especially the characteristics of biosignals are highly dependent on the person, i.e., the inter-individual variance is large. When developing context detection algorithms that utilize biosignals, data from many persons is necessary.

The volunteer was accompanied by an annotator, who used the PDA based program to mark the true context for reference purposes. A scenario was written to make sure the volunteer does the pre-defined tasks and visits the predefined locations during the measurement session. Certain level of freedom was left for the volunteer, e.g., to choose which shop or restaurant to visit.

The data collections were performed in autumn 2003 and in autumn 2004. The **scenarios** (or manuscripts) contained outdoor and indoor activities and visits to different places. Table 2 lists the scenario of 2003 data collection. In 2004 scenario the fixed tasks were followed by a 4-hour free period, during which the volunteers annotated the contexts themselves using PDA. During the free period the volunteers were free to do what they wanted as long as they carried the data collection equipment with them. The volunteers were recruited via bulletin board and electronic news ads at a local university. Table 3 summarizes the volunteer details and recording lengths of Palantir data collections.

Table 2. Scenario for Palantir data collection.

Category	Task
"Home"	Sitting at home
	Lying
	Sitting & reading newspaper
	Putting clothes on, going out
Bus	Walking to a bus stop
	Waiting for bus
	Travelling in bus
Restaurant	Walking to restaurant
	Queuing
	Eating, drinking, talking
Library	Walking to library
	Sitting in library, reading
Shop	Walking to shop
	Walking in shop, shopping
Home	Walking back home
Outdoor activities	Nordic Walking
	Running
Indoor activities	Rowing (rowing machine)
	Walking
	Bicycling (Exercise bike)
	Sitting, drinking

Table 3. Volunteer details and recording lengths of Palantir data collections.

2003 **n = 16**	**Age (yrs)**	**Weight (kg)**	**Height (cm)**	**BMI (kg/m^2)**	**Recording length (hh:mm.ss)**	**2004** **n = 12**	**Age (yrs)**	**Weight (kg)**	**Heigth (cm)**	**BMI (kg/m^2)**	**Recording Length (hh:mm.ss)**
Min	19	53	160	20,44	1:27:12	Min	19	60	167	21.5	5:49:18
Max	33	95	189	30,32	2:54:03	Max	49	85	190	26.4	7:41:39
Mean	25,81	77,50	178,75	24,14	1:57:15	Mean	27.1	76.6	179.2	23.8	6:43:49
Std	±4,34	±12,72	±7,73	±3,00	±0:22:09	Std	±9.2	±7.6	±6.2	±1.9	±0:32:38

16.6 Data Library

The collected raw data (circa 31 hours in 2003 and more than 80 h in 2004) were synchronized, calibrated, re-sampled, converted and finally gathered into a data library, consisting of the raw data (proprietary file formats, etc) and processed data. The processed files are of formats **rec**, **psv**, **mflt**, **wav** and **xml**. The **rec**, or *EDF – European Data Format* is used a lot in the field of biomedical engineering. The EDF format stores multi-channel data in binary format. The format allows the channels in one file to have different sampling rates. It also has an informative header for storing information like sampling rates, scaling information and physical units of each channel. The **PSV** format was developed in the Palantir project. It is based on the common CSV and TSV file formats. The acronym PSV stands for prefix separated values. The format is a textual format that is primarily used to represent tabular data. The **MFLT** format was also developed in the Palantir project. It is a binary format that is used in Palantir to store six channels: one channel for timestamps and five for irregularly (but simultaneously) sampled data. The **WAV** files are normal wave audio files and the accompanying **XML** file describes the start time of the audio, since WAV does not store the start time.

All the data in rec, psv, mflt and tarf formats can be viewed with Matlab and SigView (Signal Viewer). **SigView** has been written at VTT. SigView requires .NET framework to be installed on the computer. It also requires lots of memory since it reads all the signals into memory. For viewing Palantir data, 1GB of RAM is recommended. Figure depicts how SigView displays some of the data with annotations.

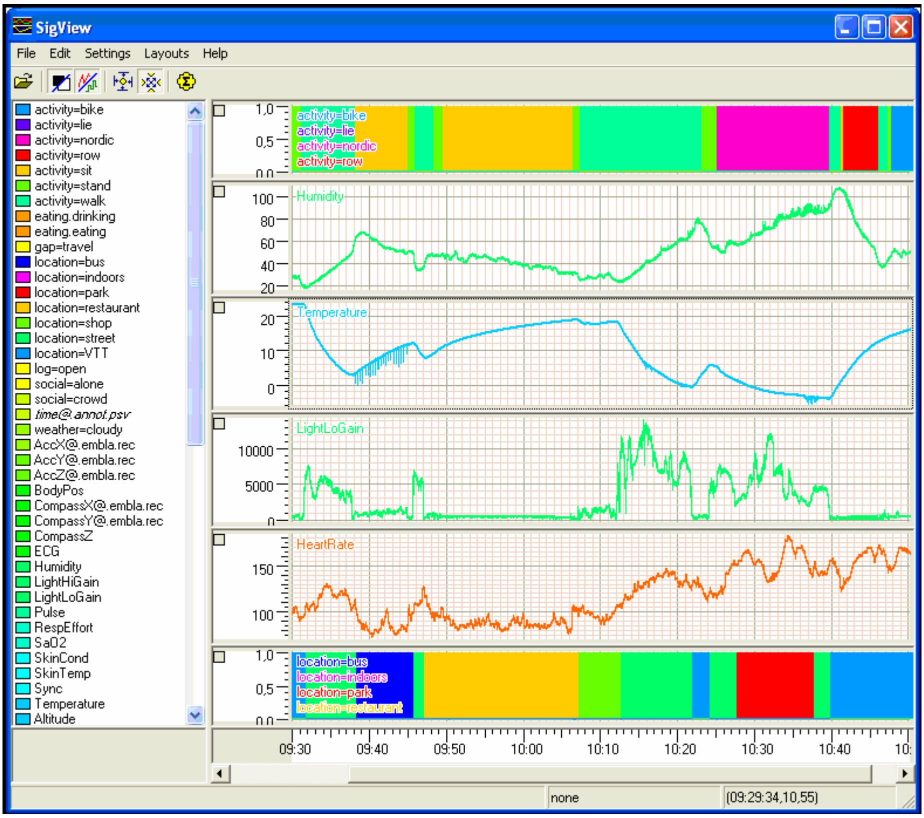


Figure 3. SigView signal viewer for viewing Palantir data.

16.7 Experiences from Data Collection

The six context types (activity, eating, gaps in life, location, social context and weather) and their context values worked reasonably well in the data collection and annotations.

The **activities** are reasonably clear to annotate. From annotation point of view, there were some difficulties with activity annotations for example during the visit to library. If the volunteer is standing and checking a book, then walking a few steps to take another book, it is practically impossible to annotate in a second-by-second basis the activities to "standing" and "walking". Instead, a third context value, e.g., "dwelling" might prove useful in such long duration measurements as in Palantir. The "dwelling" context would be useful also during transitions from one activity to another. For example, before going out, the volunteer is looking for things to take with him and constantly changing between "walking" and "standing" activities. From analysis point of view "dwelling" activity is not an exact definition, but if defined as "mixture of short duration walking and standing" it would represent the data better than either "walking" or "standing" would do.

Eating is also reasonably clear to annotate. Drinking on the other hand is very short duration activity, so its duration is only a few seconds, if annotated exactly.

"Gaps in life" context type includes in Palantir data collections two context values: "waiting" (most often for bus or for food in the restaurant) and "travelling". These are clear definitions and worked well during this data collection. It was a good idea to separate the "gaps in life" and "activity" context types. If the person receives a business call during a bus trip, we can simply deactivate the "travelling" value from "Gaps in life" context type, while the "activity" and "location" context types still reveal that the person is sitting in a bus (but no longer has a "gap" in his life).

"Location" context type had in Palantir 2003 data collection several specific context values for indoor and outdoor locations plus general "indoor" and "outdoor" context values for locations that did not fit into the more specific categories. The general context values were not used much, which shows that the more specific context values covered the locations of the measurement session very well. In longer, more independent measurement sessions, exact definition of all indoor and outdoor locations was not possible and the general values were used more.

The **social contexts** change quite rapidly if defined as the number of people around the volunteer. This is very challenging for annotation.

The **"weather"** context type was annotated mainly for background information. It was also thought that it is better to have the weather information in the annotation file than in a separate document. The weather information can be relevant, e.g., when interpreting light, temperature and humidity data. However, it was not used as an input to classifiers in data analysis.

The **annotation application** and the **PDA** (iPaq) worked well in the annotation use. The battery life was not a problem, since iPaq switches the power off if the device is not used for a while. Also powering up is fast since the last used application opens on the screen first and continues from the state it was left in. Because of the small screensize and pen-controlled UI, activating a wrong context value (by touching wrong radiobutton with the pen) was quite easy to do accidentally. Often, the context was corrected immediately by touching the correct radio button. However, a short duration, false context was stored in the log file. From this point of view, enlarging the font size would help, but from another point of view also this will cause problems. Namely, all the context names will not fit on one screen on iPaq and scrolling is necessary to find some contexts. Since all of the contexts are not visible all the time, the hidden contexts are easily forgotten and not annotated at all.

The advantage of using another person, the **annotator**, to mark the true context is that the volunteer is free to behave more naturally. Also the annotator can focus on annotating (besides surviving in traffic). The disadvantage is that use of another person as annotator is expensive, especially if even longer duration measurements are done. The presence of another person also affects how the volunteer behaves. In longer measurements, e.g., 6-hour measurements, presence of an annotator might become annoying. Thus, in longer measurements it is better to let the volunteer do the annotations himself, or to annotate offline, after the measurement session, e.g., from an audio recording. This, however, can be very time-consuming.

As a whole, the equipment worked well during the data collections. In 2003, the **Databrick PC** and its custom made battery worked well. There were no problems with the Databrick battery life even in the longest, almost 3-hour recordings. The only disadvantage of Databrick and its battery is the still reasonably big size and heavy weight. These are not a problem in normal walking, standing, etc., but, e.g., during running they do affect the way a person runs. Also **Embla** recorder worked well. There were no problems with Embla battery life either. The weakest points with Databrick and Embla were the connectors and wires. The connectors had to be taped very tightly together and enough cable length had to be reserved to ensure faultless operation of the system. During, e.g., running the equipment tends to move a lot in the rucksack, which means a lot of stress on the cables and connectors. **Suunto X6HR** worked well during the measurements. To some users (smaller and larger than average), chest belt curvature was problematic. Different sized belts for different sized persons would help here. **Garmin GPS navigator** was first used by placing it in the rucksack's right edge, just behind the volunteer's right elbow. Although the device was visible through a netted fabric, the reception was not good enough in that location. A holder was made for the GPS navigator and it was placed on top of the shoulder, attached on the rucksack strap. From the new location, the reception was very good. The GPS signal was received also in the bus if the volunteer was sitting on a window seat.

16.8 Changes Between Data Collections

Between 2003 and 2004 data collections, many changes to the data acquisition system were done. For example, the Databrick PC was no longer used as one storage media in the rucksack. The 2004 data acquisition system relied heavily on Embla, the 19-channel solid-state data recorder. Also sensors and their places were altered. For example, the piezo respiratory effort belt was replaced with RIP (Respiratory Inductance Plethysmograph) respiratory sensor and sensor box was moved from rucksack strap to hip. Annotation system still used the PDA and the annotation application for marking the true contexts. Additionally, audio and photos were stored to verify the annotation.

16.9 Context Recognition

Context recognition aims at automatic recognition of contexts from the raw signals. Thus, the key task is to develop algorithms that classify the data into interesting classes of contexts. The raw signals were first visualized with annotations. Careful visual and statistical analysis revealed, which signals contain most information on contexts. After visual analysis, **feature signals** (e.g. sliding mean, variance, etc) were calculated from the raw data. The feature signals had a sampling frequency of 1Hz. Next, a priori information was used to select features to be calculated. For example, we know that walking and running have reasonably constant frequency even between volunteers. Gait frequency in walking is typically 2 Hz, while in running it is 2.5–3 Hz.

Six of the 212 features calculated were selected to be input to classifiers. The **selected features** were 1) peak frequency of up-down chest acceleration F_{peak}(chestacc,y), 2) median of up-down chest acceleration Med(chestacc,y), 3) peak power of up-down chest acceleration P_{peak}(chestacc,y), 4) variance of back-forth chest acceleration Var(chestacc,z), 5) sum of variances of 3D wrist accelerations ∑Var(wristacc,3D), 6) power ratio of frequency bands 1–1.5 Hz and 0.2–5 Hz measured from left-right magnetometer on chest P_1(chestmagn,x).

The Palantir sensory equipment, didn't allow discrimination of sitting and standing activities from each other. Therefore sitting and standing were combined into one class, resulting in **seven target classes for activity classification**: 1) lying, 2) sitting/standing, 3) walking, 4) Nordic walking, 5) running, 6) rowing (with a rowing machine) and 7) cycling (with an exercise bike). Three different classifiers were used in classification. All of them were given the same set of six features as inputs.

For **classification**, two decision trees were applied: a custom decision tree and an automatically generated decision tree. Additionally, an artificial neural network (ANN) was used as a reference classifier. Decision trees have been successfully applied to activity recognition earlier. The custom decision tree was selected to represent simple and transparent approach based on human inference. The automatically generated decision tree was selected to see how well the automatic tree generation algorithm performs compared to human-made rules. The feature signals were given to classifiers as inputs and the classifier results were compared with PDA annotations (targets). In the comparison, each sample represents one second. The classification results are summarized in Table 4 and in Figure 4. The results show that the custom-made tree performs best in classification of most of individual activities. However, as the time spent in different activities during data collection was different, the automatically generated decision tree obtains the best overall results. This is because the algorithm optimizes overall classification performance, not the performance on as many individual activities as possible.

Table 4. Percentages of correctly classified seconds for each classifier. 1 = Custom-made decision tree, 2 = Automatically generated decision tree and 3 = Artificial Neural Network.

	1	2	3
Lie	87%	83%	74%
Row	69%	56%	59%
Exercisebike	79%	82%	75%
Sit & Stand	96%	95%	96%
Run	97%	97%	22%
Nordic	90%	72%	52%
Walk	58%	78%	79%
All	82%	86%	82%

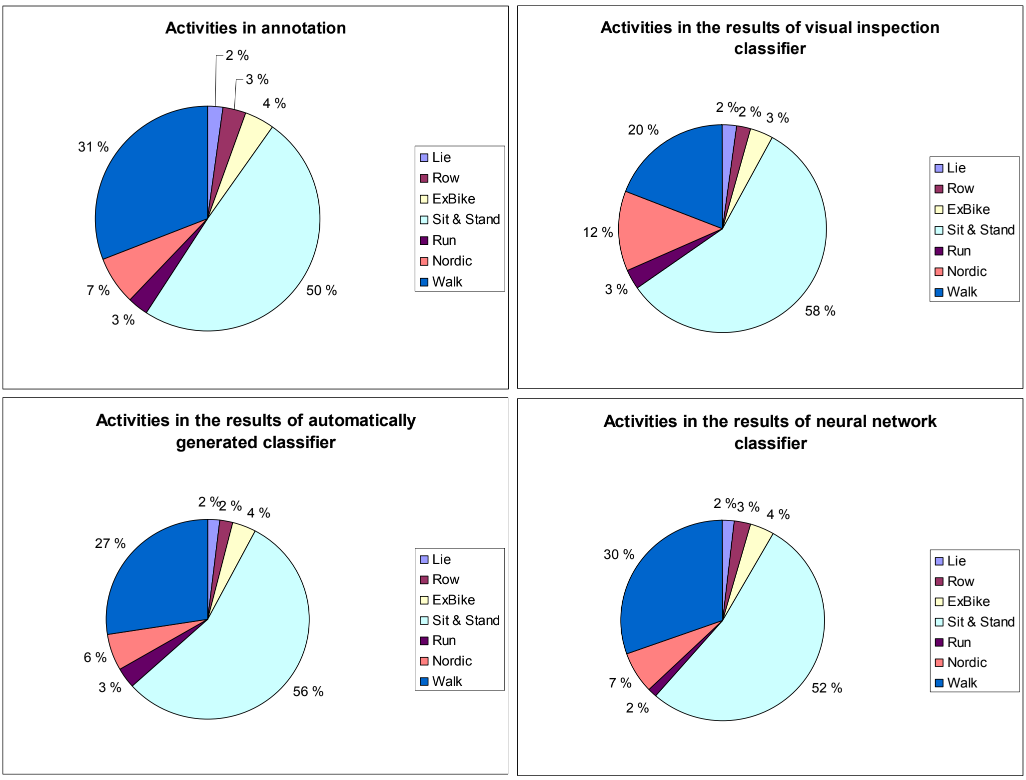


Figure 4. Portions of different activities in annotation and in classification results of the three classifiers.

Figure 4 presents the results in form of a pie. This could be, e.g., the distribution of activities during one day (now it shows distribution of activities over the whole data set).

The **activities** (running, walking, etc.) and postures (lying, standing, etc) can be recognized with sensors fairly well. Also **locations** (indoors, outdoors, vehicle) can be detected well. However, the higher level contexts like **"gaps in life".** which was defined in this data set as "traveling" and "waiting for bus" are more difficult since they would require robust detection of lower level contexts and then a combination of these. **Eating and drinking** were also difficult contexts for this data set, since they don't produce easily detectable signals on any of the collected channels. The **social context** was defined as "alone or in a crowd" in this project. Recognition of speaking was studied for recognition of social context. Results with the current algorithm are promising, but not yet mature enough for true recognition of the social context, which would be a combination of many lower level contexts like speaking.

In general, the low level contexts can be recognized with good accuracy, but the further one goes in combining the lower level contexts to form new, upper level contexts, the harder it gets to achieve good recognition accuracy. The basic postures (e.g. lying), basic activities (e.g., walking) and basic locations (e.g. indoors) can be recognized with good accuracy when using correct combination of sensors. However, recognition of more complex activities (e.g., tennis which is combination of sitting, standing, walking, running and jumping) or more accurate locations (library) requires considerably more sensors and processing. With higher level contexts, there is also the problem of defining them intuitively so that most people agree with the definitions. For example the "gaps in life" context can be problematic to define exactly. Exact definition of the context is the cornerstone of accurate annotation and thus that of accurate context recognition.

16.10 Experiences on Sensors and Signals

Especially useful sensors are those, who react fast to context changes. These were for example: 1) acceleration, 2) compass (magnetometer), 3) light (illumination), 4) GPS and 5) audio.

More difficult sensors for context recognition purposes are those who react slowly or poorly to context changes. In this study, following sensors represent this category: 1) temperature, 2) humidity, 3) heart rate and 4) respiratory effort.

Also difficult sensors for context recognition are sensors whose signal is easily disturbed by artefacts, e.g., because of movement or light. In this study, following sensors were prone to artefacts: oximeter (finger) and oximeter (forehead).

Acceleration sensors are very useful for activity/posture recognition as they react fast and robustly to changes in activity and posture. The placements of the two 3D acceleration sensors on wrist and on rucksack strap were good. The signal from rucksack strap sensor is useful for recognition of many activities, e.g., walking, running and biking with exercise bike. The signal from wrist is useful in detection of activities involving frequent hand movements, e.g., Nordic walking. Standing/sitting separation would require signal from lower parts of body.

Compass (magnetometer) sensors are not as reliable in activity detection as acceleration sensors because there are lots of metallic objects in our everyday environment. They cause small changes to Earth's magnetic field and thus artifacts to magnetometer signal. Even then, activities consisting of periodic movements can be detected reasonably well also from magnetometer signal. Such activities are for example walking and running. In analysis and calibration of magnetometer signal, it must be noted that in Finland the strongest component is perpendicular to the earth surface. Thus the sensor measures not only compass direction, but also posture of the person.

Biosignals provide useful information for exertion level of an individual, but they don't react to changes in the type of activity. For example **heart rate** (calculated from **ECG**) clearly shows the intensity an individual is performing a task, but it doesn't give any hint on whether the person is running or biking. There are also large differences between individuals, e.g., the resting heart rates are different and also changes in heart rate after certain exertion are different. In addition, the heart rate remains high for a rather long time after the exertion has been stopped, so the reaction to context change is slow. Heart rate can be measured with good reliability. However, the **respiration** signal measured with a piezo belt contains in many situations too much artefacts, e.g., due to steps. When the person is not moving, respiration rate can be detected with fairly good reliability.

Oximeter signals (**pulse, PPG** and **SaO2**) often disappear because of movement artefacts. Also external light interferes with the signals. The probe location on forehead is slightly more immune to movements, but on the other hand, it can be difficult to find a good location on the forehead to get a powerful enough signal. Finger probe provides a more powerful signal, but is more prone to movement artefacts. The pulse and SaO2 signals don't provide useful new information, but the PPG signal could be used as an estimate of relative changes in blood pressure. However, the hardware should have better attachment and cover for this to work in for mobile measurements. In this project, the signal quality was not good enough for such purposes.

Skin resistance signal reacts only to heavy sweating, not to small changes in "dry" skin resistance. After heavy sweating the material behind the sensor got wet and the signal does not return from saturated state during the rest of the measurement. Because of these hardware problems, the signal was not used in context recognition. The DC current should be replaced with AC current.

Skin temperature shows changes in skin temperature, but more clearly it shows changes in surrounding air temperature, i.e., moving between indoor/outdoor locations. Thus recognition of exercise types, etc. from skin temperature is difficult with current signals and with current equipment. The signal can be useful in stable conditions (indoors), but also this sensor reacts to changes bit too slowly. This sensor was a bit faster than the environment temperature sensor.

Body position sensor did not work in the mobile measurements. The signal provides some information on movements, i.e., whether the person is stable or not, but the acceleration signals are more useful for this purpose.

Light sensor was a useful sensor for detecting artificial light as it reacts fast to changes in illumination level. It was used for detection of indoor/outdoor. Environment sensors **humidity** and **temperature** were not very useful in this data set, because they react too slowly to context changes (temperature, more than 10 minutes). However, they do help in detection of indoor/outdoor.

Humidity sensor doesn't measure just the environment humidity, but also the humidity caused by sweating. Also, when moving from cold outdoors to warm, but dry indoors, moisture from warm indoor air concentrates on the surface of the cold sensor and the signal saturates, showing "extremely humid". The signal output corrects as the sensor warms up, but the delay is critical for context recognition purposes.

Audio can be very useful for context recognition purposes since it reacts immediately to context changes. However, interpretation of the data is difficult because of varying conditions.

Altitude and **location** (from GPS) are useful signals that react fast to changes. Altitude was not very useful for detection of the selected activities. However, speed was calculated from location data and used for detection of the "vehicle" context.

When developing a context-aware device with some sensors there are several issues that will affect the accuracy of the automatically detected context. First of all, the **selection of the sensors** is important. Not only it is necessary to select sensors that measure relevant information, but the sensors also need to have fast enough response time, broad enough measurement range and stable enough performance. If self-calibrating sensors can be used, they are surely advantageous. For example, the luxmeter used as a reference when calibrating SensorBox light sensor was calibrated simply by placing a cover on the sensor and letting the sensor self-calibrate. Even though such calibration only removes the bias, it helps a lot with sensors that easily start drifting. Similar approach could be thought also with other sensors. Second, the **measurement site** is highly important in many cases. For example acceleration sensors on chest and on wrist produce very different output depending on the activity. Thus, using many similar sensors in different places can provide new information.

16.11 Summary

Results of Palantir project as a whole are promising for activity recognition. Especially activities involving periodic movements of the body can be detected well using acceleration sensors. Also magnetometer signals can bring extra information in addition to acceleration signals. A good example of this is detection of rowing from acceleration or magnetometer signals.

“Higher level” activities that do not contain monotonic movements, but rather irregular combination of many different movements (e.g., football and other ball games, etc.) are more challenging for context recognition. Recognition of these activities could be based on a totally different approach. E.g., in Palantir, use of GPS signal was found informative for detection of football (person moves a lot, but stays within a small region).

Indoor/outdoor/vehicle recognition can also be done with good accuracy using light, humidity and temperature sensors accompanied with a GPS receiver. Placement of GPS receiver is crucial for proper reception of the GPS signal. GPS receiver on shoulder, attached to rucksack strap turned out to be a good location. Proper placement of other sensors is important as well. Placement in upper body is advantageous, because there the sensors are not easily covered by clothes or desks.

Biosignals like heart rate, respiration, skin temperature and oximeter signals are not highly useful in detection of contexts like location or type of activity, because biosignals react slowly to context changes (if they react at all). Instead, biosignals carry information on the intensity level the person is doing something. Also resting heart rates of different people are very different and since the same test induces very different changes in heart rate of different people, development of automatic algorithms is not straightforward. One approach could be to use individual training of detection algorithms (the algorithms would be adapted to the user who buys the product).

For context recognition, we used a selected set of 1Hz **feature signals** computed from the raw data. Selection of features was done using visual inspection of features against annotations and using the statistical distributions of each feature signal in different contexts. For classification we used a custom-made decision tree, automatically generated decision tree (Matlab) and an artificial neural network (multilayer perceptron). All three classifiers were given the same inputs. Use of different classifiers might improve classification results slightly, but major improvement is probably not possible. The limiting factors are as in every classification task: data quality and annotation accuracy. If the signals measured do not contain enough information on the phenomena, no classifier is able to classify the data correctly.

From sensor and device point of view, there is a growing amount of **different sensors** available that can be used in ambulatory measurements: acceleration, magnetometer, GPS, respiratory effort, heart rate, altitude, temperature, humidity, light, etc. One interesting sensor that was not used in Palantir is the **gyroscope** sensor that is suitable for posture or tilting measurements.

Developments in **solid-state memory** technology and **wireless communication technologies** will give a boost to ambulatory measurements. Solid-state memories are not sensitive to movement (unlike hard-disks) and are thus more suitable for ambulatory measurements. Wireless communication technologies (e.g., VTT SoapBox) will free the ambulatory measurements from many of the cables needed today. The biggest problem in ambulatory measurements will be **weight and size of batteries** needed.

Palantir Context Data Libraries offer possibilities for further data analysis and exploration. The analysis done during the Palantir project is by no means a complete analysis of the data libraries. The results presented in the reports represent tentative results that could be done with the resources reserved for this project. In the future, the Palantir context data library will be used in other research and development projects as well.

Publications

[P1] M. Ermes, J. Pärkkä, I. Korhonen, J. Mäntyjärvi, "Monitoring Daily Activities and Sports with Wearable Sensors." IEEE Transactions on Information Technology in Biomedicine. (Submitted)

[P2] J. Pärkkä, M. Ermes, P. Korpipää, J. Mäntyjärvi, J. Peltola, I. Korhonen, "Activity classification using realistic data from wearable sensors." IEEE Transactions on Information Technology in Biomedicine, Vol. 10, 1, Jan 2006, pp. 119–128, Doi: 10.1109/TITB.2005.856863. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1573714.

[P3] J. Pärkkä, I. Korhonen, M. Ermes, P. Korpipää, A. Ylisaukko-oja, "Kontekstintunnistus on haastavaa." Prosessori, marraskuu 2005 (Elektroniikan Suunnittelu – erikoisnumero), pp. 66–69. http://www.prosessori.fi/es05/ARKISTO/PDF/KONTEKSTINTUNNISTUS.pdf.

[P4] J. Pärkkä, L. Cluitmans, P. Korpipää, "Palantir Context Data Collection, Annotation and PSV File Format." Pervasive 2004 Workshop: Workshop towards Benchmarks and a Database for Context Recognition, Vienna, Austria, April 21–23, 2004. http://www.wearable.ethz.ch/uploads/media/WorkshopProceedingsPervasive2004.pdf.

[P5] P. Korpipää, J. Pärkkä, L. Cluitmans, "Representing features and contexts in a data library." In: Pervasive 2004 Workshop: Workshop towards Benchmarks and a Database for Context Recognition, Vienna, Austria, April 21–23, 2004. http://www.wearable.ethz.ch/uploads/media/WorkshopProceedingsPervasive2004.pdf.

[P6] M. Ermes, Context Sensing with Wearable Sensors and Data Fusion. Tampere University of Technology, 2005, MSc thesis. 69 p.

[P7] P. Korpipää, Blackboard-based software framework and tool for mobile device context awareness. PhD thesis, VTT Publications; 579. Espoo, 2005. 225 p. ISBN 951-38-6669-6; 951-38-6670-X. http://www.vtt.fi/inf/pdf/publications/2005/P579.pdf.

17. Physical Browsing for Ambient Intelligence (PB-AmI)

Heikki Ailisto, Ilkka Korhonen, Timo Tuomisto, Sanni Siltanen,
Esko Strömmer, Lauri Pohjanheimo, Jouko Hyväkkä, Pasi Välkkynen,
Arto Ylisaukko-oja and Heikki Keränen

17.1 Project Goals

17.1.1 Background

The vision of ubiquitous computing inherently includes natural interaction between humans and digital devices embedded in their environment. The desktop metaphor ([R1]) works well in the office, but it is not so well suited to ubiquitous and mobile computing ([R2, R3]). The failure of the WAP was at least partly due to the unnatural and clumsy way of selecting and browsing used.

This project studied and developed physical browsing as a means of human computer interaction by using mobile personal terminals communicating with the ambient smart devices and services. This can be seen as a building block for fulfilling the ubiquitous computing vision. Figure 1 depicts our idea.

Figure 1. Physical browsing paradigm.

Physical browsing is a very promising and potential concept. While analyzing the field of ubiquitous communication for Tekes, significant commercial and technical potential in the personal area communication applications, including physical browsing was seen ([R4]). Ideas close to physical browsing have been suggested ([R2], [R5], [R6] and [R7]). Ulmer and Ishii ([R5]) developed the idea of Phicons, which serve as physical icons for the containment, transport and manipulation of online media in an office environment. Their suggestion does not discuss the role of mobile personal devices, such as smart phones or Personal Digital Assistants (PDAs), but instead rely on fixed devices, such as digital whiteboards, projectors, and printers. Kindberg and co-workers study infrastructure to support "web presence" for the real world ([R6]), their main idea being connecting physical objects with corresponding web sites. Infra-red (IR) beacons, electronic tags or bar codes are suggested for creating the connection. We regarded the application potential much wider than accessing web pages associated with physical places or objects. After the PB-AmI project a commercial implementation of Physical Browsing, the Near Field Communication (NFC) is gaining ground (www.nfc.org). NFC is supported by Nokia, Philips, Sony, Mastercard, Visa and many other companies. VTT is a member in NFC Forum and a leader in large European project Smarttouch (www.smarttouch.org), which concentrates on development and trials of near field communication technology.

Physical browsing may be used for information retrieval related to physical objects and the environment, but it is also an essential mechanism for controlling nearby objects with a handheld device – the concept of a generic user interface requires simple and natural methods for choosing of and communicating with the nearby objects.

For object communication paradigm, the initial step of selection of the object to communicate with or to control for is of great importance from usability point of view. Several sub-paradigms may be identified: 1) ScanMe (scanning the vicinity to get a menu list of the available objetcs to the UI device, and making the selection in the virtual space), 2) PointMe (choosing the object by optical browsing), 3) TouchMe (choosing the object by touching it), and 4) NotifyMe (the terminal alarms the user according to the user profile and context). As a user interaction paradigm, these alternatives support each other and have each both advantages and disadvantages, depending on the usage situation and application.

Also the DragMe (5) paradigm can be envisaged for cyberinteraction between tagged devices ([R8]). Furthermore, in the application layer, the concept of Universal Remote Control (6) can be significantly improved when combined with object selection paradigms and application of simple tags.

The three main alternatives for implementing physical browsing in communication level are visual codes (e.g. barcodes), electromagnetic methods (e.g. RF or RFID tags) and IR (e.g. IR beacons). Each of these alternatives has its benefits and drawbacks making them suitable for certain applications. In certain cases hybrid methods may be worth consideration. In a hybrid method the connection is initiated by one method, say visual, and the bulk of communication is done by an other method, say Bluetooth RF.

17.1.2 Objectives

The main purpose was to develop, demonstrate and evaluate the physical browsing paradigm and the generic user interface paradigm to be used with mobile terminals (smart phones and PDAs) in a ubiquitous digital environment.

Different communication methods, Visual Tags, IR and RFID tags were to be used and evaluated. The goal was to analyse the available methods and their relative strengths and weaknesses compared to each other and to more traditional selection/browsing methods. We were also to propose appropriate implementation technologies and evaluate their applicability and limitations. A demonstration and trial framework were to be established based on a suitable combination of RF, IR and RFID technology to allow experimentation and (laboratory) user trials with user interaction concepts including both ScanMe, PointMe and TouchMe, and as applications information browsing and generic UI.

Using this framework, user trials with real end users and functional demonstration applications were to be carried out and usability issues related to 1) different technologies and 2) the user interaction concepts will be studied.

Physical browsing offers a natural way for human computer interaction in many applications of mobile and ubiquitous computing. Since smart phones and other mobile terminals, such as smart phones and PDAs, are becoming very common place, it is most tempting to use them as tools of interaction between the user and the ambient intelligent devices and services.

The open programmability of PDAs and new Symbian® based smart phones offers the possibility of realising physical browsing in devices existing today and those to come during the next few years. Therefore, physical browsing with mobile terminals (smart phones) has a substantial business potential.

17.2 Methods

17.2.1 Hardware Platform

The work has been mostly technology acquisition by building demonstration platforms and demonstrators. The main issue here has been the new short range technology standard NFC – Near Field Communication. The demonstrators have been used for technical evaluations, usability studies, public demonstrations and promoting non-public customer contacts. In addition to the technology acquisition, there have been some technology reviews (see also chapter 3, Results).

17.2.2 Middleware for Tag Management and RFID Based Applications

Realisation of physical browsing middleware and demo applications in a mobile device with RFID technology are described in this section.

Physical browsing can be used for multitude of applications and several demo applications were implemented during the project. In order to make the realisation dynamic and flexible in the sense that different tag-reader types can communicate with various applications, a Tag Manager middleware was designed. The middleware, reader devices and application demonstrations were implemented for Nokia 6600 mobile phone. Nokia 6600 was the most advanced model using the popular Series 60 Symbian platform available when the selection was made. It runs the Symbian operating system, which allows access to various APIs (Application Interfaces) in the C++ language allowing the exploration of various Physical Browsing scenarios.

For demonstrations, an attached RFID reader was used (Figure 2a). Due to ubiquity of barcodes a commercial barcode reader pen made by Baracoda was added to the hardware selection (Figure 2b). The pen uses Bluetooth to connect to the phone. Both the RFID reader and the barcode pen can be used simultaneously.

a) b)

Figure 2. Nokia 6600 phone with RFID reader attached (a) and a Bluetooth barcode reader pen (b).

Tag Manager middleware was designed to abstract the reader devices from the demonstration applications. The Tag Manager fulfills the following main requirements:

- Tag Manager provides an interface for tag readers, so that different tag readers, either software based or software and hardware based can be attached to the phone and possibly used simultaneously.
- Tag Manager allows multiple applications or services from different vendors to co-exist in the same phone and even use the same tags.
- Applications or services using the tags can be implemented to work without any network connections. However network based service discovery mechanisms can be implemented on top of the Tag Manager.
- The Tag Manager limits the user interface design as little as possible. Visual outlook is important in many applications, such as games, which makes replaceable selection menu component an important feature.

The Tag manager and its relation to different software and hardware components are presented in Figure 3. The components drawn with dashed lines are either not currently implemented or modified to make use of the Tag Manager. The Tag Manager is implemented as a Symbian server. It is typically loaded when one of the tag readers is started. Tag readers connect to the Tag Manager via the Symbian inter process communication (IPC) mechanism and feed the contents of the tag to the manager whenever a tag is read.

The RFID reader provides its own hardware based user interface for tag reading and no UI functionality of the phone is needed. In this case the tag reader software is implemented as a background process connected to the Tag reading hardware. Software-based readers, such as a matrix code reader utilizing the built-in camera of the phone, typically utilize the phones own user interface and can be implemented as normal applications.

Applications and services can be attached to the Tag Manager in two ways: The first is by using dynamic DLLs called TagConsumer plugins, also called starters. The second way is that applications already running connect to the Tag Manager by using the Symbian client-server approach. See references for more detailed description of the Tag Manager middleware.

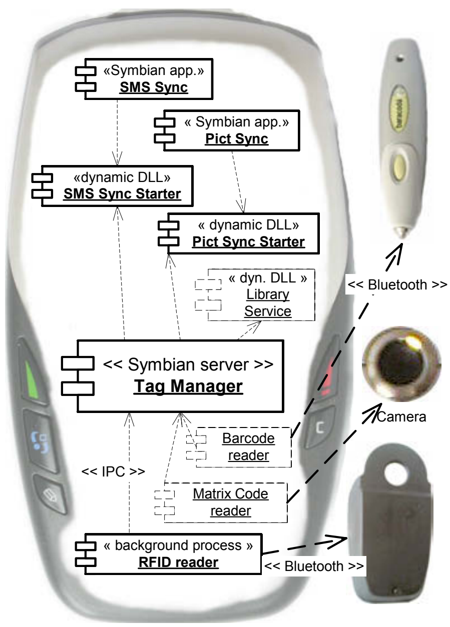


Figure 3. Deployment diagram of Tag Manager.

17.2.3 Tag Reader System to Support User Evaluation of Touching, Pointing and Scanning Paradigms

Nearby objects can be selected for communication with a mobile terminal using three alternative natural selection paradigms of

- touching objects with the mobile within reach (TouchMe),
- pointing visible objects with the mobile (PointMe),
- scanning either visible or invisible objects in the vicinity, and selecting items discovered by scanning from the terminal UI (ScanMe).

Physical Browser – a mobile terminal prototype supporting all these selection paradigms – was built for user evaluation of these different physical browsing paradigms. This prototype was also used in the MIMOSA project (Chapter 18), as the RFID sensor technology being developed in MIMOSA enables the construction of UHF RFID sensor tags which support all these physical browsing paradigms. Physical Browser was demonstrated at Pervasive 2005 ([R9]).

The physical browsing system consists of a tag reader (physical browser) and tag emulators. Both of these use SoapBoxes developed by VTT as RF communication and sensing units. The functionality of remote sensor equipped RFID tags are emulated by remote SoapBoxes and the functionality of a reader by a central SoapBox together with iPAQ 5400/5500 series PDA (Figure 4). Internet access is available by connecting the PDA via built-in WLAN card to WLAN access point.

The user interface is managed with a web browser (Internet Explorer), which displays HTML pages or is able to launch applications associated with different resource types, depending on the MIME type of the message, for example, displaying video in a multimedia player.

Figure 4. The Physical Browser components: iPAQ (right), with the embracing sleeve removed (left). The sleeve contains a central SoapBox, laser beam unit, and batteries. The SoapBox has a pointing button (shown red in the middle of the SoapBox) to activate the laser unit and the IR transmitter. The nozzle around IR led is pointing out from SoapBox.

Functional software components of the Physical browser application inside PDA are shown in Figure 5. The components are Java 1.1.8 compatible and run on Jeode[1] Java virtual machine. The PDA has a lightweight Personal HTTP server (available publicly from Acme[2]) which supports servlets. The Physical Browsing Servlet is used to construct and provide a dynamic home page: a list of available links when scanning, or when multiple tags are selected. The servlet communicates via the SoapBox driver unit with the Java serial port driver for the PocketPC. If a unique tag (containing the URL) is selected, the TagReader engine invokes the default browser of iPAQ with the specific web resource (the URL) as an argument. When presenting multiple links, the engine launches which is the web browser with the URL of the local PhysicalBrowsingServlet as an argument. The user then selects the proper link from the PDA GUI, or tries to select a tag again by touching or pointing.

---

[1] Formerly Insignia owned product, product bought by Esmertec
[2] Acme Server, http.//www.acme.com/java/software/Acme.Serve.Serve.html

Remote SoapBoxes can be programmed to send their data at regular intervals, or when a specific event occurs within the remote SoapBox. The remote SoapBox sensor board is equipped with a light sensor to detect visible light, and a proximity sensor, which consists of an IR transmitter and receiver. The IR receiver in the proximity sensor is also used to detect the pointing signal emitted by the IR LED of the central SoapBox. The remote SoapBox is programmed to regularly wake up from a sleep mode, measure proximity, and detect any IR pointing signal.

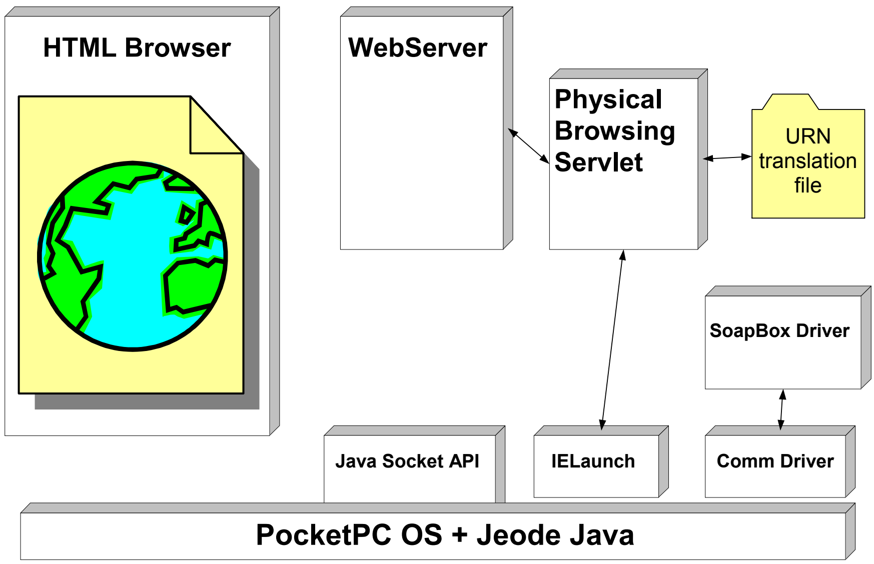


Figure 5. The Physical Browser architectural components. The personal device has an embedded HTTP server, which also supports servlets. The Physical Browsing servlet is responsible creating dynamically a HTTP page with a list of detected hyperlinks. The detection happens via SoapBoxDriver which communicates with SoapBoxes via serial port driver for PocketPC[3]. When a list is to be displayed, the servlet calls IELaunch[4] module, which activates the web browser with the local servlet URL as parameter.

---

[3] RXTX, http://www.rxtx.org/.
[4] IELaunch, http.//www.pocketpccity.com/software/pocketpc/IELaunch-2001-12-13-ce-pocketpc.html.

The physical selection paradigms – pointing, touching and scanning – are implemented with remote and central SoapBoxes as follows. For touching, the proximity sensor signal level is used to detect the proximity of objects. Whenever the measured reflected beam exceeds a certain threshold, the remote SoapBox data message is transmitted with low RF power to the central SoapBox. The range of low power transmission is about 10 cm. A flag indicating proximity is encoded in the data message. By using low transmission power, artefact detection caused by proximity of objects other than the reader can be eliminated. Pointing with IR is initiated by pressing the pointing button (see Figure 6): initially the laser pointer is activated, and serves as a visual aid only. When the pointing button is released, the IR LED pointing of the central SoapBox is activated. The IR signal is detected by the remote SoapBox, which then transmits the data message with normal RF power. A flag indicating pointing is contained in the data message. Different IR LEDs with different beam half angles ranging from +-4 degrees to +-12 degrees can be used. Extra nozzles attached to the IR LED of the central SoapBox may further be used to reduce the pointing angle. Optionally, also pointing by laser and detection by illumination sensor can be used. To mimic scanning, remote SoapBoxes work as beacons. They send regularly (programmable, typically every few seconds) their data messages to the central SoapBox. The system keeps track of available tags detected within the last 10 seconds. Scanning – presenting the dynamic home page with available tags – can be launched with one of the hardware buttons of the iPAQ.

Physical Browser user evaluations for different selection paradigms are presented in 17.3.4 "User experience".

17.2.4 Visual Tags

In the beginning of the project a literary research concerning the concept of the pointing paradigm with mobile devices was done. Moreover a deeper consideration was carried out to point out applications and situations where pointing would be a natural way of interaction between a human and the ambient world in a mobile environment. The aim was to create easy and natural access to information or services. In this subtask we identified a camera phone as suitable platform for demonstrations. This is because they are common, and most of the new mobile phones have a camera included in them. Furthermore, people carry their mobile phones with them almost all the time and they are used to them. On the whole people could use ordinary devices that they are familiar with and no additional extension parts would be needed.

In addition we identified few situations where the use of the mobile phone was difficult due to the clumsy user interface or where the access to service or information required too many steps from the user. Among these situations we selected those where the user would clearly benefit from the pointing user interface and where pointing would be a natural way of communicating with the environment and in addition where the visual tag could be used. Adding a visual tag to printed media does not bring any additional cost, thus they are especially suitable for mass media. In some other situations other type of tags could be used as well.

We identified several situations where the user would benefit from the pointing interface, see Figure 6. For example, somebody is in the bus stop and sees an interesting advertisement with an URL in the bottom of it. With some phone models typing a simple URL like http://www.google.com/ might require even 100 key presses. Would she or he be likely to start browsing it with this kind of impractical user interface? – Probably not. People are not likely to write down the URL and browse it later either. Should there be a visual tag beside the URL, the user could get the web page simply by pointing at it with the mobile phone. Other examples of practical situations for pointing user interface are: calling a person by pointing at a picture of him/her, buying a bus ticked by pointing at a tag in a bus stop, ordering something from a catalogue by pointing at a picture of the product, fetching a local map, getting a localised bus schedule, etc.

Figure 6. The user is establishing a phone call by simply pointing at a visual tag.

The general problem with mobile devices is that there exists no general standard platform, thus implementations are often device dependent. The Symbian operating system provides some kind of standard platform for certain kind of mobile phones but the number of different mobile devices and developing environments is still large. Between several choices we selected Symbian Series 60 for the development platform, which is also the platform for the Tag Manager middleware.

There is a variety of different types of visual tags: SpotCode (SpotCode), CyberCode (Rekimoto 2000), Snowflake, Data Matrix (DataMatrix), QR Code (QR Code), ArrayTag, Aztec Code, MaxiCode, ARTag (Fiala) etc. Some of them are standardised like Data Matrix and QR Code. All the available public visual tags, however, proved to be a somewhat inflexible for our needs. Thus our own visual 2D matrix code was developed. Our visual tag enabled a flexible way for e.g. to try out different information types in a matrix code.

The development of the tag reader was carried out in two parts. First the computer vision part for detecting a tag, that is image processing algorithm development, was done in Matlab environment and then implemented in Symbian environment. This approach turned out to be good for developing methods for the mobile image processing. A special focus was set to overcome the limited memory resources and processing power of the mobile device.

During the project the visual tag reader was integrated to the SW platform. The visual tag system and it various use cases were demonstrated in several public occasions.

17.3 Results

The results of the project are discussed from different viewpoints: technical sub-systems (hardware and software platform), tag technologies (IR, RFID and visual code), user experience, scientific publications and post-graduate studies, and projects originating from this theme project.

17.3.1 Hardware Platform

The following demonstration platforms have been developed:

- iPAQ integrated with a SoapBox, the latter one providing the following interaction facilities: IR-based PointMe, short-range radio communication and sensors for gesture recognition
- S60 mobile phone with an RFID reader (platform is implemented together with Nokia)
- General purpose NFC platform for interfacing sensors and other devices with NFC enabled mobile phones. The first NFC technology demonstrator (touch-based reading of a temperature sensor and a personal weight scale) was implemented with a non-commercial NFC phone obtained from Nokia.

Several demonstrators based on the platforms above have been used in usability studies, public demonstrations and non-public customer contacts.

A trial of using FEM simulations for estimating the effects of conducting materials to inductive near field links has been made. The results were close to the corresponding test results. A follow-up contract research project has been started.

17.3.2 Application Examples of Physical Browsing with RFID Technology

Demonstrations show how tags can be used to invoke actions, such as placing a phone call, performing information retrieval or mobile payment. In some cases it is viable to have a possibility of having several services available by touching one tag, instead of one service per tag. This behavior is demonstrated with SMS and Picture synchronization applications. The Tag Manager enables linking multiple services to a single tag. This makes it possible to create personalized services without the need to produce a separate tag for each user (with her preferences). This approach combines a cost effective solution for providing the tag infrastructure and personalized services.

Instant availability of the service is one key factor of the ubiquitous use of mobile services. The long startup time consisting of starting the scanners and establishing the wireless connections prevents the use of the prototype in real life use. Integration of the reader to the mobile phone should solve the startup delay problem in the prototype.

Several demonstration applications have been implemented to test the feasibility of the TouchMe paradigm of the Physical Browsing concept. In general, all examples fit well the concept idea, making the use of the applications easy and ubiquitous.

Perhaps the simplest application is the call-by-touching case. A tag attached to the business card serves as a convenient shortcut for calling. The RFID tag attached to the business card contains the phone number and is used by the caller application to launch the phone call. Also pictures of friends and relatives could be used for phone call initiators.

Two applications with similar characteristics were built to evaluate the TouchMe paradigm in device to device communication: SMS Sync application sends all new SMS messages from the phone to the PC when a RFID tag attached to the PC is touched (Figure 7). Similarly the Picture Sync application sends all new photos taken with the phone built-in camera to the PC. Both applications use the same RFID tag which provides a test for the Tag Manager’s service selection properties described in section 2.2: if the phone is in the basic state, an application selection list is presented and if one of those two applications are visible to the user the tag is passed to this application.

The RFID tag attached to the PC contains the unique Bluetooth address of the PC Bluetooth adapter. The standard Bluetooth Object Exchange (OBEX) protocol is used for transmission; hence no special software for the PC is needed. Both applications also keep track which messages or pictures are sent to which computer.

Figure 7. SMS messages are sent to the PC by touching the RFID tag.

Mobile payment was also explored with a mock up application. A vending machine with an option to order soda by calling a certain number was equipped with an RFID tag (Figure 8.a). The tag contained a phone number of the service. Using the tag with the TouchMe paradigm removed the need to punch in the phone number and open the call. The tag-augmented vending machine was perceived to be easy to use without the additional hassle of looking up and keying in the phone number.

The barcode reader was also tested alongside the RFID reader. A barcode containing the ISBN (International Standard Book Number) number of a book was read with the Baracoda barcode pen (Figure 8.b).

Figure 8a. Soft drink is ordered by touching RFID tag.

Figure 8b. ISBN barcode in a book offers a link to a university library database.

The scanning launches the factory installed XHTML browser of the phone and gives it a URL containing a query to a local university library database. A page containing book availability information in the library is returned.

17.3.3 Physical Browsing with Visual Tags

One of the results was creating our own 2D visual marker – VisualTag, see Figure 9. The marker consists of black and white squares, three black edges and one varying black and white edge. The number of data cells of the marker is not fixed. We have used sizes from 6x6 to 30x30 i.e. from 36 to 900 data cells. The smallest size is more than enough for a simple marker ID (256 different). These can be used for applications where there is no need to read more information from the marker and the small physical size of the marker is essential. Currently, our markers can store up to 600 bits of information which means about 70 uncompressed characters or e.g. a capacity for 180 digits.

Figure 9. VisualTags. On the left there is a marker that consists of 36 data cells and a marker on the right has 18x18 i.e. 324 data cells.

The orientation of the marker is free as it is detected from the discontinuous edge. The data consists of three parts: an information type or application identification part, an information part and an error correction and detection part. The information can be of any type: text, special characters, numbers, binary data etc. The information or the application type is an identification number (ID) and it tells to the detection application how to react to the data. We have currently defined for example the following ID types: text, URL, phone number, SMS, control, marker-ID and game. When our implementation detects a URL type marker it opens the www-browser in the phone with the address coded in the information part. Respectively, a phone call is established automatically in the case of a phone number. A control tag is used for controlling settings of a mobile device. The text is shown to the user etc.

The developed mobile image processing methods for the visual tag system were implemented in several software packets. The VisualTag Generator is used for generating visual tags. Two versions of VisualTag Generator were produced; one version was implemented using Java Servlet technique. It includes a graphical user interface and it is accessible via the Internet using normal web browser. Another stand alone version implemented for the cases where the network access is not available.

The mobile phones we used in our development work are capable of capturing still images up to a resolution of 640 x 480 pixels (VGA). The capturing is a relatively slow process and it takes nearly one second to capture one image with that resolution. The phones are also able to capture low resolution images of size 160 x 120 pixels (QQVGA) with more than ten images per second. The capturing behaviour was the reason why two versions of the VisualTag reader were implemented. The first one uses only low resolution images. It is able to function very fast but it demands a bigger physical size of a tag. The other one uses low resolution images for smooth aiming and for predicting the tag existence probability. A VGA resolution image is captured only when a certain probability threshold level is exceeded. The actual information is decoded from the VGA resolution image. This version can detect physically very small tags as can be seen in the following screenshots, Figure 10.

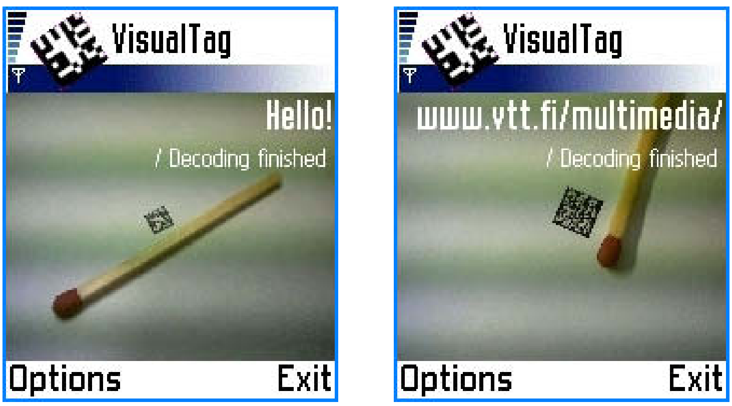


Figure 10. Decoding the Visual Tag content.

The screenshots were taken automatically immediately after decoding. On the left there are 100 data cells in the area of 4 mm x 4 mm. On the right there is a tag of size 6 mm x 6mm with 324 data cells. The number of data cells is more than enough for text 'Hello!' and web address 'www.vtt.fi/multimedia/', respectively. In these two cases the phone, Nokia 3660, was equipped with an add-on lens. With the lens the camera is able to focus near the marker. The reading distance with the lens is about 6–8 cm.

The next two examples in Figure 11 show detection from different view points. The tag is of size 24 mm x 24 mm. The screenshot on the left were taken using VGA mode and the one on the right using low resolution mode. The add-on lens was not used in these cases.

Figure 11. Detection and decoding from skewed viewpoint.

The recognition speed depends on how many other objects are situated near a tag. The low resolution mode might take up to 200 ms to finish detection if the image is full of small objects. The detection is typically completed in 30–80 ms in normal cases. The VGA mode completes detection procedure typically in 800–950 ms (including 600–700 ms capture time). These timings were measured using Nokia 7650 and 3660 which have a CPU running at 104 MHz. Current mobile phones are faster and more efficient but their actual timings were not measured. Both of the implementations are binary compatible with several phone models. VisualTag runs successfully at least on the following Nokia phones: 7650, 3650, 3660, 6600, 7610, 6670, 6680 and 6630.

VisualTag reader was implemented also as a dynamic link library for Symbian. All the needed image capturing and processing methods and decoding algorithms are included in the DLL. It is a simple process to implement new visual tag applications using the DLL.

The usability of VisualTag was tested successfully with a group of elderly people who normally have great difficulties to use mobile phone's tiny keypad. The group used VisualTag for making phone calls and for sending short messages to e.g. relatives.

17.3.4 User Experience

Pbysical Browser was used to evaluate touching vs. pointing, and some parameters to these paradigms ([P11]). In the following methods and results are shortly introduced. The test situation is shown in Figure 12.

Figure 12. A user pointing a remote tag. A test tag embedded in medical product package is visible on the table.

In the evaluations the users performed several tag selection tasks by both pointing and touching with different configurations:

- Touching: A) selection by pressing a button and B) selection without button.

- Pointing: A) selection with laser only, B) selection with IR only and C) selection with laser and IR combined.

Users were asked to use different selection techniques to select tags being at different distances. The evaluation was made when the user was sitting, standing or walking. After the selection tasks, the users were asked to score the different touching and pointing configurations on six usability aspects on a questionnaire.

Touching was perceived as an interaction method easy to learn and use. The most salient difference between the two touching configurations was in effortlessness. Touching with button was perceived to require more effort than without the button. Based on these results, it seems that there is no clear reason to choose or reject either of the touching configurations so far. We suggest that as both of the configurations are easy to implement, the final decision should be made by the user.

Touching without a button should however be a distinct mode that can be easily set on and off to prevent continuous unintentional selections, for instance, when the reader device was in a bag in contact with a number of tagged objects.

Of the three pointing configurations, pointing with only IR beam was the least preferred, which probably largely reflected the real unreliability of the prototype. Only circa 90% of IR beam selections succeeded with no certain technical reason for the unsuccessful ones. The laser-only configuration required accurate aiming and thus it was evaluated more negatively than the combined IR and laser. Pointing with the combination of IR and laser was the most preferred. These results show that pointing with a mobile device should provide visual feedback to the user, but in addition to this pointing should have some insensitivity to aiming errors.

There are also other options to implement pointing with visual feedback and insensitivity to aiming errors than a system of two overlapping beams. The laser beam could be widened to reduce the effect of shaking and to allow easier aiming. Widening the beam would however lower the intensity of it. A lower intensity beam is not as visible to human eye in bright sunlight as a narrower higher intensity beam, and detecting it requires sensors that are more sensitive.

Our results show that pointing is a useful selection method and the users prefer pointing to touching if the distance to the tag is more than 1–1.1 m. A pointable tag has to have a range that significantly exceeds 1.1 m to be usable at all, and it is not useful to calibrate the default pointing distance (if the technology needs it) to be less than one metre. Furthermore, social acceptance issues may be relieved with different interaction methods. Earlier studies have brought out that touching tags with a mobile device in public places could be socially embarrassing for the user. Pointing is more unnoticeable and thus perhaps more acceptable in public places, until the interaction methods become common.

We will see an increasing amount of tagged objects in our environment. With non-visual tags there is a problem of visual discovery of available information. Currently Bluetooth devices and NFC devices use their own technology logos. Moreover many local marketing solutions based on Bluetooth have adopted their own logos. Thus the logos market only the technologies. To avoid chaos, we suggest that the physical selection methods as an abstraction could be shown in the tags, together with an icon or symbol to represent the type of service ([R10]).

17.3.5 Media Coverage

The results of the PB-AmI have enjoyed good media coverage. We organized a press conference in Tampere in October 2004 resulting in national TV, radio and press coverage. Some applications, for example with the elderly users received attention in TV news and in the press.

17.3.6 Projects Originating from PB-AmI

PB-AmI gave VTT an excellent position to enter the emerging Near Field Communication development in 2005–2006. VTT is leading a large ITEA project SmartTouch which focuses on developing physical browsing technology and applications. Near field communication technology has also been piloted in Tekes project WSENSE (Long Term Wellness Monitoring by Wireless Low Power Sensors at Real Life Settings). Furthermore, a new Tekes project on visual appearance of the tags is to be started.

17.4 Conclusion

The general goal of the Physical Browsing for Ambient Intelligence project was set to research and advance the use of physical browsing user interface paradigm. This new paradigm was seen as one solution to the problem of how to bridge physical and digital worlds. The concrete goals of the project were to implement the technical components needed for physical browsing realisations with different tag technologies; to study and enhance usability and user experience and advance the development of physical browsing concept by spawning new R&D activity and applications.

The project resulted to the implementation of hardware platforms necessary for the experiments, a dynamic middleware supporting heterogeneous tag types and applications, i.e. the Tag Manager, software tool for generating and reading Visual Tags even in challenging conditions and a flexible test set-up for user studies. The user studies produced interesting results about usability and user experience which showed that the users welcome the new user paradigm.

Publications

[P1] H. Ailisto, L. Pohjanheimo, P. Välkkynen, E. Strömmer, T. Tuomisto, I. Korhonen, "Bridging the Physical and Virtual Worlds by Local Connectivity Based Physical Selection." Personal and Ubiquitous Computing Journal. Pers Ubiquit Comput, 2006. DOI 10.1007/s00779-005-0057-0.

[P2] L. Pohjanheimo, H. Keränen, H. Ailisto, "Implementing TouchMe paradigm with a mobile phone Smart Objects and Ambient Intelligence." sOc-EUSAI. Grenoble, 12–14 Oct. 2005. Smart Objects and Ambient Intelligence A Joint Conference Proceedings (2005). Pp. 87–92.

[P3] H. Ailisto, L. Pohjanheimo, H. Keränen, P. Alahuhta, "Physical Browsing – a novel HCI paradigm for people on the move." In: Embedded Real Time Software ERTS2006, Toulouse, France, Proc of ERTS2006, CD only. 2006.

[P4] H. Keränen, L. Pohjanheimo, H. Ailisto, "Tag manager: a mobile phone platform for physical selection services." IEEE International Conference on Pervasive Services 2005 (ICPS'05). Santorini 11 Jul. 2005 (2005). Pp. 405–412.

[P5] L. Pohjanheimo, H. Ailisto, J. Plomp, "User experiment with physical pointing for accessing services with a mobile device." EUSAI Workshop on Ambient Intelligent Technologies for Wellbeing at Home. Nov. 2004.

[P6] M. Sipilä, P. Raatikainen, M. Åberg, I. Norros, L. Norros, A. Mämmelä, T. Frantti, J. Holmberg, H. Ailisto, A. Lattunen, J. Latvakoski, S. Horsmanheimo, "Topics in 4G networks and hardware." In: Proceedings of the International Workshop on Optical and Electronic Device Technology for Access Network. Aalborg, 2005. Aalborg University, Keio University. Tokyo, Japan, 2005. Pp. 5–10.

[P7] T. Tuomisto, P. Välkkynen, A. Ylisaukko-oja, "RFID Tag reader system emulator to support touching, pointing and scanning." Pervasive Computing. München, May 2005. In: Advances in Pervasive Computing. Austrian Computer Society. Wien, 2005. Pp. 85–88.

[P8] P. Välkkynen, T. Tuomisto, "Physical browsing research." Proceedings of the Workshop Pervasive Mobile Interaction Devices. München, May 2005. Enrico Rukzio, Media Informatics Group, University of Munich. Wien, 2005. Pp. 35–38.

[P9] P. Välkkynen, I. Korhonen, J. Plomp, T. Tuomisto, L. Cluimans, H. Ailisto, H. Seppä, "A user interaction paradigm for physical browsing and near-object control based on tags." In: Proceedings of Physical Interaction Workshop on Real World User Interfaces, Udine, IT, 8–11 Sept., 2003. University of Udine, HCI Lab., Department of Mathematics and Computer Science, 2003. Pp. 31–34.

[P10] S. Siltanen, J. Hyvakka, "Implementing a Natural User Interface for Camera Phones Using Visual Tags." In: Proc. Seventh Australasian User Interface Conference (AUIC2006), Hobart, Australia. CRPIT, 50. Piekarski, W. (Ed.). ACS, 2006. Pp. 113–116.

[P11] P. Välkkynen, M. Niemelä, T. Tuomisto, "Evaluating Touching and Pointing with a Mobile Terminal for Physical Browsing." Proceedings of NordiCHI 2006 Conference, October 2006, Oslo.

[P12] H. Ailisto, J. Plomp, L. Pohjanheimo, E. Strömmer, "A physical selection paradigm for ubiquitous computing." In: 1st European Symposium on Ambient Intelligence (EUSAI 2003). Veldhoven, NL, 3–4 Nov. 2003. Ambient Intelligence. Lecture Notes in Computer Science Vol. 2875. Aarts, Emile et al. (Eds.). Springer-Verlag. Berlin, 2003. Pp. 372–383.

[P13] H. Ailisto, I. Korhonen, J. Plomp, L. Pohjanheimo, E. Strömmer, "Realising physical selection for mobile devices." Physical interaction (PI03) workshop on real world user interfaces, 2003. Pp. 38–41.

[P14] A. Ylisaukko-oja, "Perspectives for RFID in mobile phone ambient intelligence applications." ID WORLD, Barcelona, 18 Nov. 2004. 16 p.

[P15] M. Hillukkala. "NFC-yhteyden toteuttaminen sulautetuissa laitteistoissa." Oulun yliopisto, sähkö- ja tietotekniikan osasto. Diplomityö, 2006. 40 p.

References

[R1] D. Smith, "Designing the Star User interface." Byte, April 1982, pp. 242–282.

[R2] R. Want, M. Weiser, E. Mynatt, "Activating everyday objects in Darpa/NIST." Smart Spaces Workshop (Gaithersburg, MA, July 1998), USC Information Sciences Institute, 1998. Pp. 140–143.

[R3] A. Dey, “Understanding and Using Context.” Personal and Ubiquitous Computing (2001), 5, pp. 4–7.

[R4] H. Ailisto, A. Kotila, E. Strömmer, “Ubicom Applications and Technologies.” TEKES, 2003. 44 p.

[R5] B. Ullmer, H. Ishii, D. Glas, “mediaBlocks: Physical Containers, Transports, and Controls for Online Media.” In: Proceedings of SIGGRAPH ‘98 (Orlando, FL, July 1998), ACM Press. Pp. 379–386.

[R6] T. Kindberg, et al. 2000. “People, Places, Things: Web Presence for Real World.” In: IEEE Workshop on Mobile Computing Systems and Applications WMCSA’00 (Monterey CA, December 2000), IEEE Press. Pp. 19–28.

[R7] S. Pradhan, “Semantic Location.” Personal Technologies (2000), 4, pp. 213–216.

[R8] J. Rekimoto, Y. Ayatsuka, “CyberCode: designing augmented reality environments with visual tags.” In: Proc. Designing Augmented Reality Environments (DARE 2000). Pp. 1–10.

[R9] T. Tuomisto, P. Välkkynen, A. Ylisaukko-oja, “RFID Tag reader system emulator to support touching, pointing and scanning.” Pervasive Computing. München, May 2005. In: Advances in Pervasive Computing. Austrian Computer Society. Wien, 2005. Pp. 85–88.

[R10] P. Välkkynen, T. Tuomisto, I. Korhonen, “Suggestions for visualising physical hyperlinks.” In: Proc. of Pervasive Mobile Interaction Devices Workshop, Dublin, Ireland, May 2006.

18. Microsystem Platform for Mobile Services and Applications (MIMOSA)

Aarne Oja and Eija Kaasinen

18.1 Background

MIMOSA was an integrated project in the EC 6th framework programme (IST-2002-507045). The project advocated a mobile-phone centric realization of ambient intelligence. The consortium included major industrial partners and leading European research institutes. The Finnish partners of MIMOSA were VTT, Nokia, Suunto, and MAS. The duration of the project was from January 1, 2004 until June 30, 2006. The other industrial partners of MIMOSA were STMicroelectronics, Legrand, ÅMIC, Cardiplus, Sonion. In addition to VTT, the participating research institutes and universities were CEA-LETI, CSEM, EPFL, FhG-ISIT and CNRS-LAAS.

18.1 Objectives

In the MIMOSA vision, the personal mobile phone was chosen as the trusted intelligent user interface to Ambient Intelligence and a gateway between the sensors, the network of sensors, the public network and the Internet. MIMOSA developed an open technology platform for implementing ambient intelligence in different application areas. The well defined platform allows a fast and focused development of both basic technology solutions as well as system-level applications and services. MIMOSA focused to develop micro- and nanosystems in several areas of the MIMOSA open platform.

In the environment domain, *i.e.*, in an intelligent environment of a mobile user, MIMOSA contributed by developing wireless remote-powered sensors which exploit RFID technology, and novel low-power radios which use RF MEMS technology.

In the user domain, MIMOSA contributed by developing microsystem-based intuitive user interfaces which made efficient use of short-range communication and various sensors embedded both in the mobile handset and the local environment. MEMS based user-activity and physiological sensors were used as user interfaces to realize new applications.

In the phone domain, MIMOSA developed readers/writers for RFID tags and sensors which were closely integrated with the Bluetooth radio by sharing same RF circuit blocks. In addition, MEMS-based inertial, magnetic and audio sensors were realized and embedded in the handset, to enable novel intuitive user interfaces.

18.2 Work Breakdown and Methodology

The organization of the MIMOSA project was optimized to support the creation of an open technology and application platform which is well adapted to generating useful services for future intelligent environments. The MIMOSA work packages are represented in Fig. 1. Three work packages were led by Finnish partners, WP1 and WP3A by VTT, and WP2 by Nokia.

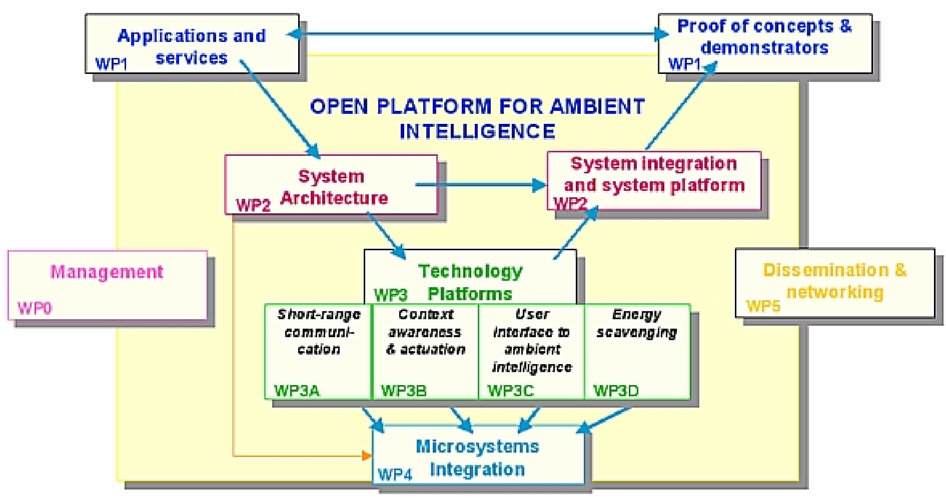


Figure 1. MIMOSA work packages.

Human-centred design in the context of the design of a universal platform and related technologies is quite extraordinary; usually human users are thought about only when individual applications are being designed onto the platform. However, many forthcoming features of applications are actually fixed by the solutions done on the platform level, and that is why VTT in MIMOSA project aimed to apply human-centred design principles in designing technical infrastructures for Ambient Intelligence solutions.

In the development of MIMOSA technological building blocks, extensive use was made of advanced IC technology and various MEMS technologies:

- Silicon MEMS technology for inertial and magnetic sensors, and for low-power, MEMS-based reference oscillators.
- Thin film AlN technology for SMR-BAW filters and oscillators for low-power radios.
- Polymer technology for physiological sensors and some novel applications.
- Advanced 130 nm SOI-IC technology for low-power RFID tags and sensors.

The development of technological building blocks made use of the usual tools that are used for developing microsystems and integrated circuit. These include various circuit simulators both in the device and system level, electromagnetic simulators for antenna design, FEM simulators for MEMS design, process simulators for IC and MEMS fabrication.

18.3 Results

18.3.1 User and Application Requirements for Ambient Intelligence Solutions

The aim of MIMOSA work package 1 (WP1) was to ensure that the development of the MIMOSA core technology is based on the needs of the users, and that the resulting technologies will be easy-to-use, useful and acceptable from the end user's point of view, as well as applicable from the application developers point of view. VTT has coordinated and carried out most of the work in WP1.

In the beginning of the project, WP1 coordinated the definition of a common vision for the project, and illustration of the vision as usage scenarios. During the first project year, the scenarios were analysed and evaluated for credibility, acceptability and technical feasibility by potential end-users, field experts, and technology and application developers in the project.

During the second year, key parts of the usage scenarios were implemented as proof of concept technology demonstrators that illustrate the look and feel of applications that MIMOSA technology will facilitate in the selected application areas (sports, fitness, health care, housing, everyday). These proof of concepts were evaluated with users and application field experts. The evaluation results and the potential of MIMOSA technologies were assessed with the technology and application developer partners in the project.

It turned out that human-centred design can be applied also when designing technical infrastructures but novel methods are needed. Scenarios and proof of concepts based on existing technologies illustrated to potential users and application field experts as well as to the project group itself what the project was targeting to. User evaluations helped in refining the scenarios. The hardest part of the work was identifying the implication of the scenarios and user feedback to the technical infrastructures. Our approach was based on identifying common patterns, and analyzing in details the interactions with users, terminals, tags and sensors within those patterns. Potential implications were discussed through with technology experts to clarify the technical feasibility of different solutions. Further work is needed to study how user feedback could be obtained more effectively and precisely, and earlier in the project.

18.3.2 RFID-Based Solutions for Short Range Communication: Tags and Sensors

The basic solution for the MIMOSA local connectivity was to make use of the 2.4 GHz ISM band. The goal was to develop a reconfigurable RF front end which can be used as the RF front end for the existing Bluetooth radio or, alternatively, as the RF front end of an RFID interrogator.

The central features of two RFID communication protocols were chosen for implementation, namely the ISO 18000-4 standard and a modification of the Palomar protocol for the 2.4 GHz band.

The primary goal was to develop passive tags and sensors, *i.e.*, tags that are powered remotely by the interrogator. The efficiency of the wireless power transfer and the power consumption of the tag circuit then became major design issues. The effective apertures of the tag and the reader antennas are restricted by their small geometric sizes. The efficiency of the RF rectification of the tag circuit is important. Two different rectifiers were designed and fabricated. One was based on CMOS-transistors on 130 nm SOI IC technology of STMicroelectronics. The other approach made use of Schottky diodes on thick SOI wafers.

The analog circuit blocks such as a demodulator, regulator, limiter, modulator, and power-on-reset circuit were also designed on the 130 nm SOI IC technology of STMicroelectronics. The digital blocks of the RFID protocol management were implemented by using a FPGA circuit. These features were later to be included in the integrated circuit.

A low-power interface circuit with A/D conversion was designed for capacitive sensors. The RFID protocol was defined to include sensor data inquiries.

The first wireless remotely-powered sensors of the MIMOSA have been integrated and tested. Figure 2 illustrates an RFID tag on a foil-type antenna. Figure 3 shows the calculated radiation pattern.

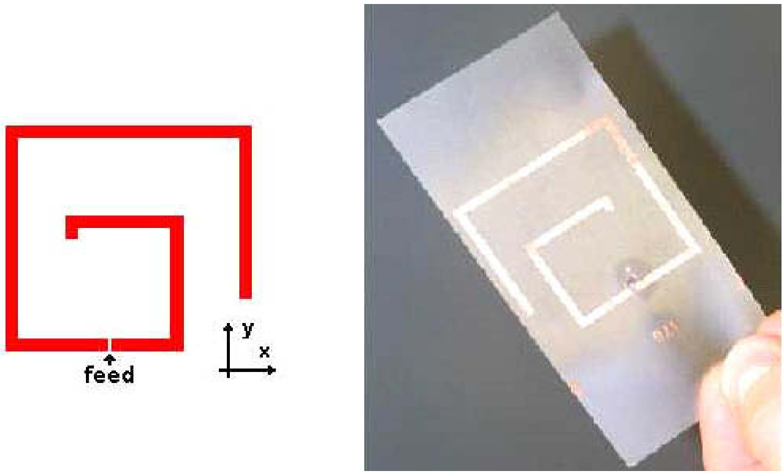


Figure 2. Foil-type antenna for WRPS.

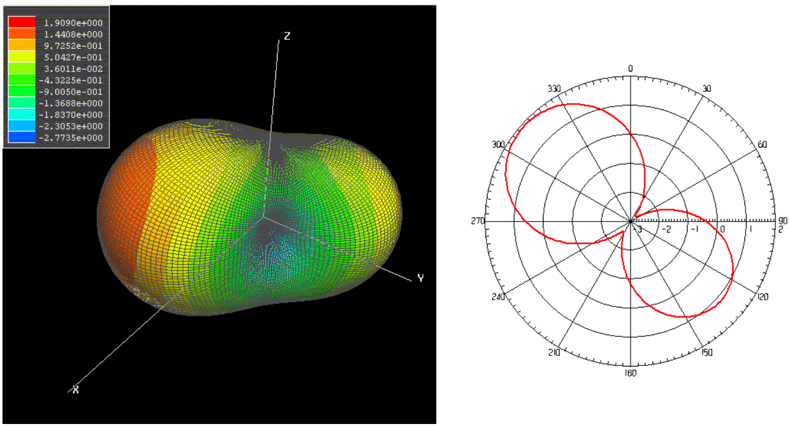


Figure 3. Radiation pattern of the antenna of Fig. 2.

18.3.3 RFID-Based Solutions for Short Range Communication: Reader

Figure 4 represents the architecture of the RFID reader RF parts which are designed to be finally integrated in a mobile phone. The major difference between a radio and an RFID reader is that the RFID reader simultaneously sends and receives signal at the same frequency. A good isolation between the transmit and receive paths is essential, otherwise the receiver low noise amplifier (LNA) gets easily saturated, especially when the same antenna is being used for both signals. Additional problems are caused by environmental reflections of the transmitted microwave signal. These reflections can amount to some -20 dB signal at the LNA with respect to the transmitted power. Environmental reflections change in time while the antenna and the nearby reflecting objects move with respect to each other. Thus the isolation should adapt to the changing environment. Figure 4 illustrates one solution to this problem. A separate compensation channel is used to superpose an appropriate signal at the LNA input to keep the LNA within its dynamic range. The amplitude and the phase of the compensation signal is controlled by a baseband logic. The circuit has now been integrated with a mobile phone platform. Tests have confirmed the basic functionality in the RFID reader and Bluetooth modes. Further tests are being made on the radio channel stability. This demonstrator will be used to communicate with the RFID sensors developed in MIMOSA.

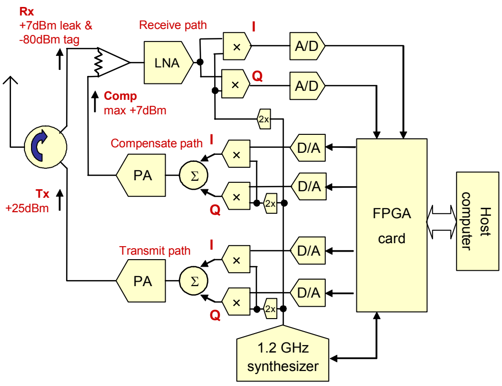


Figure 4. Block diagram of the combined 2.4 GHz RFID reader and a Bluetooth radio.

18.3.4 MEMS-Based RF Circuit Blocks

MEMS technology offers new solutions for several components in RF electronics. The general advantages of RF MEMS include low-power operation, miniaturization, small parasitics, tunability and/or reconfigurability, potential for highly integrated modules and subsystems. MIMOSA developed several RF MEMS components:

- MEMS based reference oscillators
- Thin-film AlN components for SMR-BAW filters and oscillators
- AlN-piezoactuated RF-MEMS switches
- Packaging and interconnection techonologies for RF-MEMS components.

MAS and VTT are developing monolithically integrated MEMS based reference oscillators. The device is shown in Figure 5. The operation of the device is based on a square-shaped silicon structure which is suspended above the substrate by spring structures at the corners of the square. In addition, the middle point of the square is fixed to the substrate. Other parts of the square are released from the substrate by using a special etching technique developed by VTT, *i.e.,* the so-called plug-up method. This etching technique makes it possible to release patterned MEMS structures from the substrate and yet preserve a wafer surface that can undergo the normal processing steps that are needed for fabrication of CMOS circuits.

The device that is being developed by MAS and VTT is a monolithically integrated 13 MHz reference oscillator. The functional blocks of the device include the MEMS resonator, the loop amplifier, a charge pump for the high bias voltage of the MEMS resonator, phase-locked-loop for a 2.4 GHz voltage-controlled oscillator, and a temperature sensor and associated control electronics for thermal compensation.

The first set of devices has been fabricated at VTT. Tests show that integration of loop amplifiers with MEMS resonators was successful. The next manufacturing run will narrow the gap between the sense and drive electrodes of the square plate. This should increase the electromechanical coupling to the level needed for a closed-loop oscillation. The devices have so far been tested under a pumped low-pressure environment. Wafer-level low-pressure packaging is eventually needed.

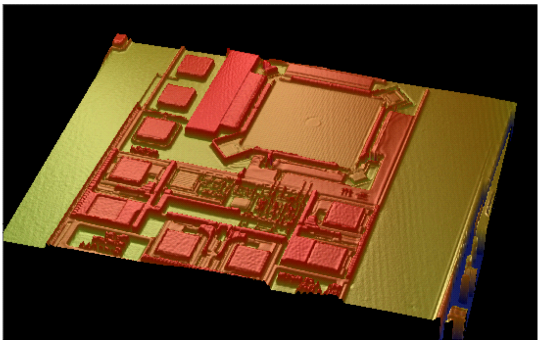


Figure 5. Optical profilometer image of a 13 MHz oscillator circuit based on a MEMS resonator.

18.4 Commercial Impact

Industrial partners will exploit the results of MIMOSA within their own applications fields. These include Nokia (telecommunications), STMicroelectronics (IC circuits and microsystems), Legrand (home automation), Suunto (sports and fitness), ÅMIC (fabrication of polymer-based microsystems), Cardiplus (tele medicine), MAS (IC circuits), and Sonion (MEMS-based microphones).

Dissemination and Publications

[P1] "STS Forum", Kyoto, Japan, 14 – 16 November, 2004. Presentation by Heikki Seppä.

[P2] "Nordic Semiconductor Meeting." Oslo in August 2005 (NSM21). Invited talk by Aarne Oja.

[P3] "MIMOSA Workshop on Invisible Electronics for Ambient Intelligence Applications." Grenoble, September 2005. Presentation by Eija Kaasinen

[P4] "ELMO Results Promotion." Helsinki, November 2005. Presentation by Aarne Oja.

[P5] E. Kaasinen, K. Rentto, L. Norros, 2005. "Ihmisen ja jokapaikan tietotekniikan vuorovaikutus." [User interaction with ambient inteligence]. In: A. Kasvio, T. Inkinen, H. Liikala (eds.). Tietoyhteiskunta. Myytit ja todellisuus. [Information Society, mythes and reality9. Tampere University Press. Pp. 347–366. (In Finnish)

[P6] V. Ikonen, J. Leikas, H. Strömberg, "Nomadic elderly: Design requirements for technologically enhanced travelling experience." Paper in 5th Conference on Gerontechnology. Nagoya Congress Center, Japan. May 2005.

[P7] P. Välkkynen, T. Tuomisto, "Physical browsing research. Proceedings of the Workshop Pervasive Mobile Interaction Devices." München, May 2005. Enrico Rukzio, Media Informatics Group, University of Munich. Pp. 35–38.

[P8] V. Ikonen, "Scenarios in ubiquitous computing system design; user-driven usages."Conference paper in HCI – 11th International Conference on Human-Computer Interaction, HCII2005. Caesars Palace, Las Vegas, Nevada, New Jersey. July 2005.

[P9] V. Ikonen, H. Strömberg, "Perspectives to evaluations of usage scenarios in intercultural projects." Presentation in 1st International Conference on Usability and Internationalization. Ceasars Palace, Las Vegas, Nevada. July 2005.

[P10] V. Ikonen, "Scenarios for intelligent sports applications." Poster in Ambience'05, Tampere, Finland, Poster in Ambience 2005 conference. Tampere, Finland. October 2005.

[P11] E. Kaasinen, T. Tuomisto, P. Välkkynen, "Ambient Functionality: Use Cases." Smart objects & Ambient Intelligence Conference. Grenoble, France. October 2005.

[P12] V. Kaajakari, J. K. Koskinen, T. Mattila, "Phase noise in capacitively coupled micromechanical oscillators." IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 52, No. 12, Dec. 2005, pp. 2322–2331.

19. BIOSEC

Heikki Ailisto, Mikko Lindholm, Satu-Marja Mäkelä,
Elena Vildjiounaite and Jani Mäntyjärvi

BioSec was an Integrated Project, co-funded by the VI Framework Programme of the European Commission, started in December 2003 and ended in November 2005 with the purpose of consolidating and progressing the European research on biometric technologies. The project had more than 20 partners from nine European countries. The project was led by Telefonica I&D, Spain, while the work package leaders were VTT, Ministry of Interior (Finland), Telefonica, Etra and University of Carlos III (Spain) and University of Cologne. The project was initiated by Finnish partners and VTT did the preparation work. Other Finnish partners besides VTT and the Ministry of Interior were Finnair and University of Tampere.

Biometric technologies provide automated means to identify a subject according to physiological or psychological features. These technologies improve the verification of the individuals and have become the key technology for securing privacy. Nevertheless, this technology is neither completely consolidated nor accessible with the required features in security, trust and confidence to the public nowadays.

The results of the work of VTT in the project fall into five areas:

1. Developing, implementing and experimenting with the biometric access control pilot in Finnair office building at Helsinki Airport. The pilot had two stages, first a fingerprint based system was used with approximately 50 users and then a multimodal biometric system with three dimensional face and fingerprint recognition was set up. In the first case VTT was responsible for the biometrics technology and implementation while in the latter case it was a joint effort between VTT, Siemens, ITI and VCON. Finnair hosted all the cases and the Ministry of Interior gave specifications.

2. Noise cancelling front end for a voice recognition system was developed in cooperation with Telefonica, Polytechnic University of Madrid and Polytechnic University of Catalonia. Telefonica developed the actual voice

recognition system to which the noise cancelling front end was integrated. The voice recognition system was demonstrated with remote and home applications.

3. VTT developed and demonstrated the concept of soft biometrics, where features such as weight, height or even body fat percentage were used to differentiate between individuals in a small or medium size group (up to 50 persons). The concept is quite novel and it suitable for cases where privacy needs override high security, i.e. in comfort or usability driven applications.

4. Accelerometer based gait recognition was developed by VTT's research group as an unobtrusive means of verifying the identity of a person carrying a mobile device, such as a mobile phone. The recognition performance was on the level of practical face and voice recognition systems. The media coverage on the invention was quite wide, for example BBC, Reuters and many newspapers carried the news and the inventors were interviewed by Japanese and South African radio stations in addition to national news media.

5. On multimodal and cascaded biometrics VTT developed together with Spanish partners Support Vector Machine and Multilayer Perceptron based multimodal biometric decision making algorithms. VTT also developed cascaded decision making engine for less obtrusive user verification

Publications

[P1] H. Ailisto, E. Vildjiounaite, M. Lindholm, S.-M. Mäkelä, J. Peltola, "Soft biometrics – combining body weight and fat measurements with fingerprint biometrics." Pattern Recognition Letters. Vol. 27 (2006), pp. 325–334.

[P2] H. Ailisto, M. Lindholm, P. Tikkanen, 2003. "A review of fingerprint image enhancement methods." International journal of image and graphics. World scientific publishing, Vol. 3, 3, pp. 401–424.

[P3] H. Ailisto, M. Lindholm, J. Mäntyjärvi, E. Vildjiounaite, S.-M. Mäkelä, 2005. "Identifying people from gait pattern with accelerometers." Proceedings of SPIE. Vol. 5779. Biometric Technology for Human Identification II. Anil K. Jain & Nalini K. Ratha (Eds.). SPIE. Pp. 7–14.

[P4] J. Mäntyjärvi, M. Lindholm, E. Vildjiounaite, S.-M. Mäkelä, H. Ailisto, 2005. "Identifying users of portable devices from gait pattern with accelerometers." Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP '05, March 18–23, 2005, Vol. 2. Pp. 973–976.

[P5] E. Vildjiounaite, S.-M. Mäkelä, M. Lindholm, R. Riihimäki, V. Kyllönen, H. Ailisto, 2006. "Unobtrusive Multimodal Biometrics for Ensuring Privacy and Information Security with Personal Devices." 4th International Conference on Pervasive Computing, 2006. Dublin, Ireland, 8–9 May 2006.

[P6] H. Ailisto, M. Lindholm, S.-M. Mäkelä, E. Vildjiounaite, 2004. "Unobtrusive user identification with light biometrics." NordiCHI 2004. Proceedings of the Third Nordic Conference on Human-Computer Interaction. Tampere, 23–27 Oct. 2004. ACM Press. Pp. 327–330.

[P7] E. Vildjiounaite, V. Kyllönen, H. Ailisto, "Empirical Evaluation of Combining Unobtrusiveness and Security Requirements in Multimodal Biometric Systems." Submitted to Image and Vision Computing Journal, Special Issue on Multimodal Biometrics.

20. WWI Ambient Networks

Jukka Mäkelä, Kostas Pentikousis, Pekka Pääkkönen, Vesa Kyllönen, Jari Korva, Mikko Majanen, Ville Typpö, Jari Tenhunen, Pekka Koskela, Jouni Heikkinen and Marko Jurvansuu

20.1 Introduction

The future wireless world will be filled by a multitude of user devices and wireless technologies. Simple-to-use, anytime-anywhere network access affordable for everyone is an undisputed must for the eSociety and i2010 where the user will expect a rich set of communication services independently of the access used. Devising technologies for true Plug-and-Play networking, efficient use of all infrastructure investments and access competition are the key technical challenges to achieve this vision.

The Ambient Networks project is addressing these challenges by developing innovative mobile network solutions for increased competition and cooperation in an environment with a multitude of access technologies, network operators and business actors. With more than 40 partners, the project offers a complete, coherent wireless network solution based on dynamic composition of networks that provide access to any network through the instant establishment of inter-network agreements. The concept offers common control functions to a wide range of different applications and access technologies, enabling the integrated, scalable and transparent control of network capabilities.

The project has three phases. Phase 1 (2004–2005) has established the overall approach and developed innovative technical concepts. Currently ongoing phase 2 (2006–2007) will prove their viability through implementation, integration, measurements and performance evaluation, enabling concurrent standardisation. The planned effort of approximately 2200 person-months reflects the strategic and technical ambitions of Ambient Networks during its second phase. The commercial viability is validated through field trials in planned Phase 3.

The figure below shows the different Planned Phases and work areas of the project.

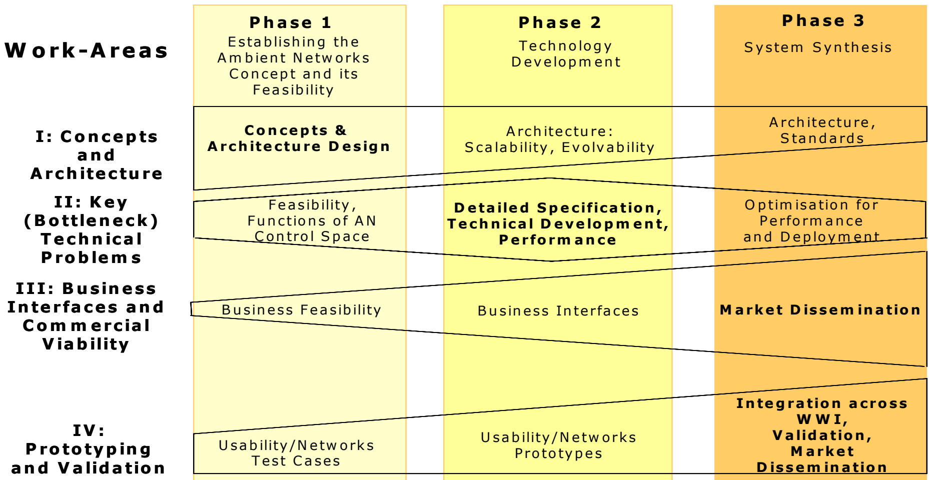


Figure 1. Project phases and work areas.

Ambient Networks brings together a strong industry-led consortium of the leading operators, vendors, SMEs and research organisations, with the determination, skills and critical mass to create cross-industry consensus for mobile and wireless systems beyond 3G. The results will facilitate the incremental market introduction of new services, and stimulate sustainable growth in the European mobile communications sector.

Currently ongoing Ambient Networks Phase 2 is a joint project of 44 partners ranging from large industrial companies to operators and SMEs together with universities and research centres around Europe. Project will continue the work done in Phase 1 (2004–2005). The project co-ordinator is Ericsson.

In Phase2 VTT participates in two of eight work packages WP-B “Mobile Connectivity” and WP-H “Ambient Networks Realization” in Ambient Networks Phase 2 project. VTT has also a task leadership in both work packages. WP-B works on mobility and moving networks. WP-H has the overall responsibility to develop the Ambient Networks Prototype. Most of the conceptual work is done in the WP-B which is validated by real implementations and prototyping work hosted by WP-H. Detailed work package descriptions can be found in [P1].

20.2 Results

A summary of mechanisms and prototypes developed by VTT is given in the following (more detailed information about the mechanisms can be found later in this chapter):

- A mechanism for session handover was designed and implemented. A notification of invention was made about the mechanism and currently the investigation and evaluation of novelty of the mechanism and possible actions to bring this research to business are in progress.
- New mechanisms for Moving Network Support were defined with a triggering mechanism to support mobility in heterogeneous networks. A prototype implementation of the Triggering Engine is available. The Triggering Engine has been notified also in the Standardisation forums (IEEE 802.21).
- Several simulation models and implementations for NS-2 network simulator were developed. The developed models can be used also in other projects if necessary.
- A mechanism for the dynamic creation of a Personal Area Network (PAN) cluster was designed.
- A gateway selection framework for moving networks was designed.
- A demonstration with the first real implementations of the all designed concepts was built.
- Implemented Prototypes were presented in several public events: WWI symposium, IST Mobile Summit 2005, PIMRC'05 Conference, CityWLAN seminar and IST Mobile Summit 2006.
- All works are described in public deliverables which are available at the project website www.ambient-networks.org.

20.3 Detailed Information about Mechanisms

This chapter gives some details about the mechanism that VTT has been developing in the project.

20.3.1 Triggering

Mobile devices are now capable of running demanding network applications, may have multiple network interfaces and, thus, connectivity options. Nevertheless, state-of-the-art mobile protocol stacks can only handle a small set of event notifications, typically related to radio access network (RAN) connectivity, user mobility, and load balancing. For example, signal strength deterioration generally leads to a base station handover (HO) in cellular voice; 2G/3G mobile phones typically opt for 3G connectivity when the user moves into a new area; and, sustained high data traffic loads may force the Universal Terrestrial Radio Access (UTRA) transport function to reallocate resources (and even perform a HO) in WCDMA 3G/UMTS networks. We argue that we need a framework to handle a much larger set of notifications caused by events that originate not only from the lower layers of the protocol stack (physical, data link, and network), as in the examples, but from the upper layers (session, transport, and application) as well.

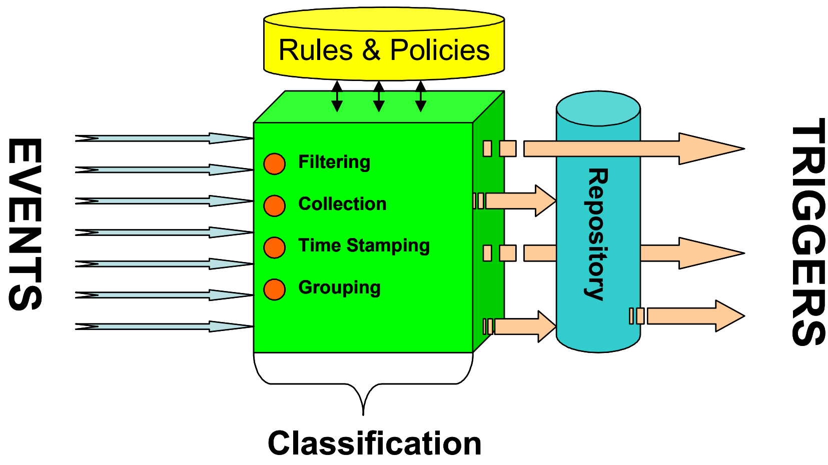


Figure 2. TRG FE Overview.

TRG is shorthand for “Triggering”, the Functional Entity (FE) which receives “events” from other FEs, processes them, and generates “triggers”, which are then sent in a standardized format to other FEs. TRG is mostly concerned with mobility related events, and any other information that can assist handover decision making. It thus has to be very compact and perhaps trade some flexibility in order to deal with events expeditiously. This is why TRG has an internal repository for triggers and delivers them in a single format. Figure 2 presents the high-level view of the TRG FE.

TRG emphasizes that all triggers coming out of the TRG FE are in a single, standard format that the consumers of triggers can rely upon. It is the job of TRG to take, perhaps, "raw" information (say, CPU load, or WiFi bandwidth utilization measurements) and deliver standardized notifications. These "standardized notifications" are always of type "ANTrigger" regardless of whether the trigger consumer receives a steady stream of CPU load measurements as they occur, an aggregate/mean value every so often, or a single notification when the CPU load has exceeded a certain threshold. This approach lends itself well to developing a standard API for receiving triggers, which can then be used by any AN FE.

20.3.2 Moving Network Support

In brief, the Moving Network Support (MNS) is responsible for the functions of network mobility management. This includes creating and maintaining Routing Groups (RGs) and selecting and maintaining gateways and mobile routers.

In Ambient Networks, a Routing Group can be formed by a group of network nodes that are physically near each other and possibly share similar movement patterns. The RG framework supports multi-hop clusters where nodes may have a special cluster-head or gateway role.

The simplest benefit of RGs is that they can provide connectivity for nodes that would otherwise be unreachable. The RG structure can, however, enable further optimizations, related to both routing and mobility management. The algorithm chosen for forming RGs has three phases: neighbor discovery, cluster-head selection and maintenance phase. Every node running the algorithm is periodically sending Hello messages and listening for incoming ones. During neighbor discovery, the nodes identify their stable neighbors within one-hop radius by assigning each neighbor a stability value. The value is updated based on associativity (received Hello messages) and link quality (signal-to-noise ratio). Those neighbors whose stability value exceeds a pre-defined threshold are marked as stable.

After discovering these stable links a cluster-head, a controlling master-node, is selected. The selection algorithm is a distributed process where each node calculates a suitability value for itself and broadcasts the value to the other stable nodes. Finally, the node with the highest value is chosen as the cluster-head. In the maintenance phase, after the RG is formed, it is possible to manage the configuration of the RG by performing operations either on single nodes or whole RGs. That is, single nodes may join or leave an existing RG, and two groups can be merged or one RG can be split into two.

To get external and global connectivity, the RG nodes need to communicate via selected gateway (GW) or mobile router (MR) nodes. To support identification, management and selection of GW and MR nodes, a Gateway Selection Architecture (GSA) has been specified. In addition to ordinary RG, GW and MR nodes, the architecture introduces special Gateway Selector (GWS) nodes. In the same way as the cluster-head nodes manage RGs, GSW nodes manage the discovery and selection process of the GW and MR nodes.

The gateway discovery process has three phases. First, the GWS node is elected during RG formation. Typically, the chosen cluster-head can be assigned also the role of GWS. Next, the potential GW or MR nodes pro-actively advertise their presence and capabilities to the chosen GWS. Finally in the third phase, whenever RG nodes needs the GW service, i.e. communication with nodes outside the RG, they send a request to the GWS, which selects the most suitable GW (or MR) for the requesting node, according to a number of parameters related to topology.

20.4 Multi Dimensional Mobility

Mobility may take place in different dimensions, which are independent of each other. We have identified seven dimensions, which in most cases can be considered as orthogonal to each other.

- Physical location: Communication endpoint moves within the same radio access technology (traditional mobility),

- Access Technology: Communication endpoint moves from one radio access technology (RAT) to another (e.g. vertical handover),
- Address Space: Communication endpoint moves between networks/devices, which use different address space (e.g. IPv4 or IPv6),
- Security domain: security context (generally speaking) has to be transferred from one to an another point in order to maintain the security properties,
- Provider Domain: Communication endpoint moves between networks/devices operated/owned by a different provider (e.g. roaming),
- Device Properties: Communication endpoint moves from a device to another, hence the system properties of the host device may change dramatically (e.g. inter-device handover),
- Time: Communication endpoint does not move spatially, but on-going communication is suspended for a while and resumed afterwards (e.g. if a user wants to pause the connection for a while, or for a temporary loss of connectivity).

Some of the trigger events relate only on a single mobility dimension, while others may require mobility actions to be performed in several dimensions. One could imagine any moving entity to have a coordinate in each dimension. Figure 3 illustrates four of these dimensions.

Whenever a movement in certain dimension takes place, the respective coordinate changes. On abstract level, mobility management mechanisms can be seen as functions updating one or more of these coordinates.

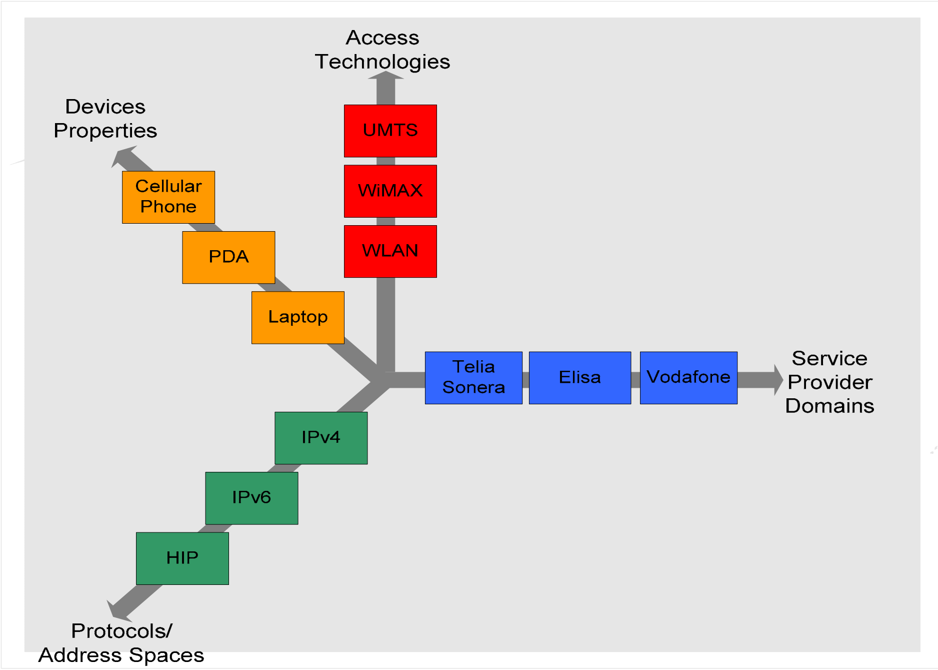


Figure 3. Orthogonality of the Mobility Dimensions illustrated with four dimension.

20.5 Prototype & Demonstration of the Concepts with Session Handover Mechanism

The scenario of the experimental study embraces three different parties:

- A content provider. It is a service provider which offers a streaming service to its customers.
- Mobile users. Different mobile users that are willing to access the content provided by the previous provider. In addition, they have the capability of using the different AN mobility mechanisms that were described before.
- Gateway nodes. These are special nodes that are able to connect the mobile users to an infrastructure network; furthermore, they are enhanced with the AN mobility procedures.

Furthermore, the scenario will encompass the typical devices that are likely to be part of the forthcoming personal area networks, such as PDA and laptops. In short, the study can be described as follows: some users meet with a friend, a RG is automatically formed, and a GW node, which allows the users to access the content provided by the video server, is selected. One of the users is using both a laptop and a PDA, and using the hint provided by an attached sensor box, the video stream shifts (session handover) from one device to the other. Later on, a new potential GW appears. It is offering a better connection, so the RG changes its current GW without any service disruption. The next section will depict all the steps within this demonstration flow, mapping them onto the concepts that were described before.

In particular, the study will aim at assessing the feasibility of the following challenges:

- Dynamic formation and maintenance of RG.
- Selection and maintenance of GW.
- Dynamic change of GW, without service disruption.
- Demonstration of the Triggering FA, being able to collect, Later (according to predefined policies) and execute the corresponding actions.
- Assess the feasibility of session handover, as a consequence of some of the triggering events.

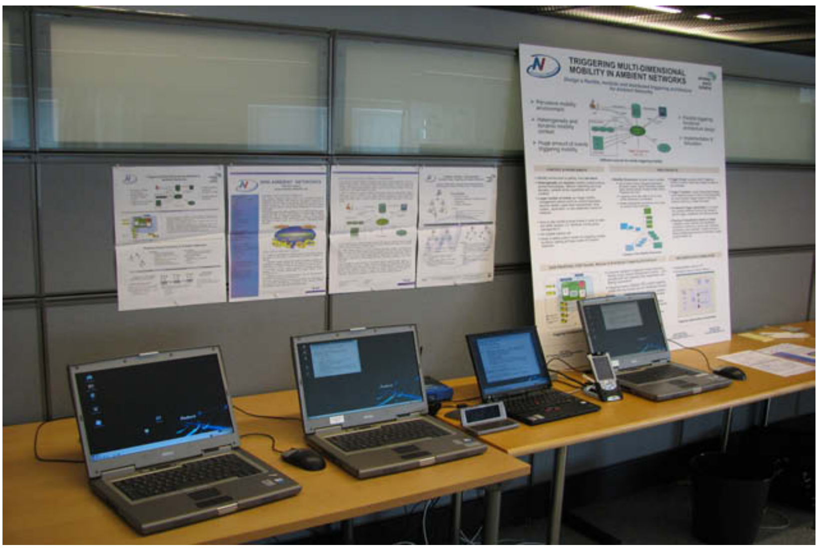


Figure 4. Testbed at VTT.

20.6 Description of the Demonstration Flow

At the beginning, the different nodes (two laptops and a PDA) are switched on. They begin with the distributed interchange of Hello messages, from which they are able to assess the associativity (which, to a certain extent, is mapped on the degree of connectivity). This information is then used to create an RG, in which one of the nodes takes the role of a cluster-head. In addition, one of the devices, which is able to connect to external networks, is selected as GW and therefore, all the other devices use it to access service provider (i.e. the video server); a streaming session is then started from the server to one of the devices. While the video session is ongoing, a new GW appears and joins the RG; according to a set of configurable policies (e.g. price, security, QoS,...), it is offering a better connection than the existing one, so all the devices (RG members) change their GW. The video session that was ongoing continues through the new GW without any service disruption, that is, the end user does not perceive any halt while watching the video. After a while, the second GW is switched off and the former one is selected again; this also takes place without the user noticing it (i.e. there is no disruption).

Afterwards, the user with the PDA wants to receive the video with it, rather than keep on using the laptop. He moves the PDA to put the screen towards him/her and the sensor box attached to it detects this change in device orientation, sending the corresponding event to the Triggering FA, which takes the appropriate steps to perform the session handover. As happened before, there is not any noticeable service disruption in the meanwhile. After a while, the user may decide to continue watching the video with the laptop, so he/she puts the PDA downwards and the stream is automatically shifted again to the laptop. Furthermore, an additional trigger can be delivered by other means, such as pushing a button in the PDA, which sends the video stream to the other laptop, thus mimicking an action by means of which the user is able to share his/her video with a friend (another device which is also part of the RG).

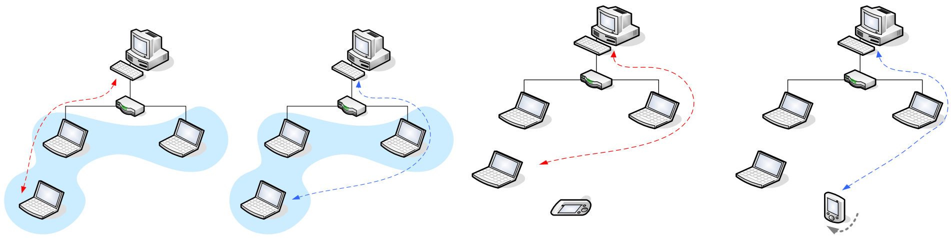


Figure 5. Illustration of the demonstration.

Publications

[P1] J. Mäkelä, R. Aguero, J. Tenhunen, V. Kyllönen, J. Choque, L. "Munoz, Paving the Way for Future Mobility Mechanisms: A Testbed for Mobility Triggering & Moving Network Support." Proceedings of 2nd International IEEE/Create-Net Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, Tridentcom 2006, Barcelona, Spain, March 1–3, 2006. IEEE/Create-Net.

[P2] J. Mäkelä et al. "Liikkuvuus saa uusia ulottuvuuksia." Prosessori (2006) No. 1, pp. 46–49.

[P3] P. Pöyhönen, R. Campos, N. Akhtar, C. Kappler, P. Pääkkönen, D. Zhou, "DEEP – A Generic Name Resolution Protocol for Heterogeneous Networks." In: 2nd IEEE International Conference On Information & Communication Technologies: From Theory To Applications (ICTTA'06), April 24–26, 2006, Damascus, Syria.

[P4] P. Pääkkönen, N. Akhtar, R. Campos, C. Kappler, P. Pöyhönen, D. Zhou "Scalability of Name Resolution for Ambient Networks." WWIC 2006, Bern, Switzerland.

[P5] R. Campos, N. Akhtar, C. Kappler, P. Pääkkönen, P. Pöyhönen, D. Zhou, "On the evaluation of the EGIST Protocol." (Published in IST Summit 2006)

[P6] J. Mäkelä, R. Aguero, D. Hollos, J. Choque," Ambient Mobility in Practice – Mechanisms for Moving Network & Triggering." Proceedings of 15th Mobile and Wireless Summit 2006. 4–5 June 2006, Myconos, Greece IST Summit 2006.

[P7] N. Akhtar, R. Campos, C. Kappler, P. Pääkkönen, P. Pöyhönen, D. Zhou "GANS: A Signalling Framework for Dynamic Interworking between Heterogeneous Networks." In: VTC Fall 2006, Montreal, Canada.

[P8] R. Agüero, J. Eisl, V. Typpö, S. Uno, "Analysis of Mobility Control Functions in Ambient Networks, 3G 2005." In: 6th International Conference on 3G & Beyond, November 2005, The IEE, Savoy Place, London, UK.

[P9] J. Tenhunen, V. Typpö, M. Jurvansuu, "Stability-Based Multi-Hop Clustering Protocol." In: Proceeding of the 16th IEEE annual International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC 2005), September 2005.

[P10] M. Eyrich, M. Majanen, E. Perera, R. Toenjes, R. Boreli, T. Leinmueller, "GSA: An Architecture for Optimising Gateway Selection in Dynamic Routing Groups." In: Proceeding of the 16th IEEE annual International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC 2005), September 2005.

[P11] J. Puttonen, G. Fekete, J. Mäkelä, T. Hämäläinen, J Narikka, "Using Link Layer Information for Improving Vertical Handovers." In: Proceeding of the 16th IEEE annual International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC 2005), September 2005.

[P12] V. Typpö, O. Gonsa, T. Rinta-aho, B. Tharon, S. Bonjour, "Triggering Multi-dimensional Mobility in Ambient Networks." In Proceedings of the IST Mobile & Wireless Communications Summit 2005, June 2005, Dresden.

[P13] A. Surtees, R. Aguero, J. Tenhunen, M. Rossi, D. Hollos, “Routing Group Formation in Ambient Networks.” In: Proceedings of the IST Mobile & Wireless Communications Summit 2005, June 2005, Dresden.

[P14] V. Typpö, J. Eisl, J. Höller, R. Agüero, H. Karl, “Research challenges in mobility and moving networks; an ambient networks view.” Proceedings of Workshops on Broadband Satellite Communications and Challenges of Mobility in conjunction with IFIP World Computing Congress. Toulouse, 22–27 Aug. 2004. Kluwer (2004).

[P15] O. Gonsa, T. Toniatti, V. Typpö, F. Meago, “Triggering and handover management concepts for mobility management in B3G networks, B3G-SA Report. Beyond 3G System Architectures Cluster.” Communication and Networking Technologies, Information Society Technologies (2004), pp. 80–89.

[P16] P. Koskela et al, “Scenarios for Network Composition in Ambient Networks: a new paradigm for Internetworking.” Wireless World Research Forum (WWRF) 11th meeting, 10–11, June 2004.

[P17] J. Mäkelä, K. Pentikousis, “Trigger Management Mechanisms for Ambient Networks.” International Symposium on Wireless Pervasive Computing 2007. (Accepted)

[P18] J. Eisl, J. Mäkelä, R. A. Calvo, S. Uno, “Ambient Networks Book Chapter.” To be published by WILEY, March 2007. (Accepted)

[P19] K. Pentikousis, R. Giaffreda, E. Hepworth, R. Aguero, A. Galis, “An Information Service Infrastructure for Dynamic Networks.” ICCGI 2007. (Submitted)

[P20] M. Majanen, K. Pentikousis, “An Evaluation of the Ambient Networks Gateway Selection Architecture.” ICWMC 2007. (Submitted)

21. Nomadic Media

Johan Plomp and Jani Mäntyjärvi

21.1 Executive Summary

The Nomadic Media project aims to find solutions for the design process of future ubiquitous media services and techniques needed for users to interact with them. Nomadic users will interact with media services in their homes and outside. They will want to use and discuss about content of media with other people wherever they are through the most appropriate available interaction modalities and devices. The issues addressed by this project are:

- Methodologies for requirements capturing and user-centred design (UCD) of situated services
- Technologies for determining and utilising the context for service selection and adaptation
- Interaction technologies allowing for personalised adaptive multimodal interaction
- An architecture for the implementation of adaptive situational services.

21.2 Background of the Project

Levels of consumer acceptance of networked services will be determined by the quality of interaction people enjoy when using their devices, the services and the content. Thus, the overall goal of the Nomadic Media project is:

> *Find solutions that allow consumers to enjoy their content and use interactive services at the times and in the places they prefer, using the devices that best suit their circumstances.*

21.3 Objectives

The ambition of Nomadic Media is to reduce restrictions imposed by platforms and devices when consumers decide on ‘the what’ (content and/or services), ‘the where’ and ‘the when’. We also aim at systems that adapt more readily to consumer’s personal preferences and needs that are enjoyable to use, and provide low entry thresholds for all sections of society.

The Finnish partners will particularly aim to find solutions to the following challenges:

- Novel user centred design methodologies for situated services.

  Development of situated services includes many stakeholders and involves necessary interaction between a large number of separate components. The success of services is difficult to predict due to the lack of comparable existing services. Furthermore, the evaluation of pilot services is challenging due to their inherent operation in everyday environments, which is very different from any laboratory environment. Early integration of users in the design process and improved business models and life-cycle management of the development of services will be needed.

- Facing the challenge of distributed adaptive user interfaces

  Interaction with applications or services embedded in the environment will happen through a diversity of terminals and modalities. Therefore the definition of a situated service inevitably includes the description of a user interface in a manner that allows adaptation to the terminals and modalities at hand and most suitable for the interaction at that time and place. The interaction may furthermore be dependent on the context and the user (profile). Accommodating these new requirements for user interface development, distribution and adaptation will be a major challenge.

- Context description and exchange

  Describing context so that it can be used also by applications and services that have not themselves produced the context information is a challenging task. An appropriate context description language is needed for establishing the user's context, enhancing media with relevant context information and selecting and/or adapting media and services according to the context. One needs to agree on a format and ontology to describe the context as derived from sensors and other relevant sources. Also methods are needed to establish to context through reasoning on perceived context cues and matching perceived context with available contextual information and services. Furthermore, exchange of context information between applications and services will require a proper platform and protocol that should be open to extensions.

- Multi-network, multi-terminal, multi-channel, multi-modal, multi-user

  Facing these "multi-" issues is a significant challenge for the development of new services. The partners in this proposal will each face new requirements due to the character of the situated services to be developed. Networks will change especially when moving from outdoor to indoor environments. Terminals range from large screens to small mobile phones, each with their own established solutions and channels for service access. New modalities will enhance the interaction styles. Furthermore, applications at home may simultaneously be operated by several people and collective services in e.g. a ski resort will allow for public and group communication. The terminals need to be used by a range of different users.

21.4 Resources

– VTT Electronics focused on research on distributed adaptive user interfaces, context description and exchange and integration of developed technology components.

– VTT Information Technology focused on research on user centred design methodologies for situated services.

– The project started in July 2003 and ended in June 2005.

– The project was managed by Johan Plomp and Jani Mäntyjärvi.

21.5 Main Results

The project developed situated applications and technologies, which were demonstrated in an airport environment:

- Bluetooth based indoor positioning and location context producer
- Context ontology for enabling airport services
- Symbian application for providing social grouping and recommendations
- Sensing equipment for ski-slope activity recognition (wearable-wireless sensing system).

Novel technology for human-computer interaction was also developed, it was demonstrated in airport and home environments:

- Gesture control device containing 3D accelerometer, 3D gyroscope and 3D magnetometer, laser assisted IR-based fysical pointing with tactile feedback
- Gesture control of public space displays
- Functional gesture control mock-up device to facilitate rapid application prototyping.

Research methodology for user centred design, e.g. methods for visualising end-user scenarios in cost efficient manner.

Visiting researcher Juha Kela stayed one year at Berkeley, California, USA.

Publications

[P1] J. Mäntyjärvi, K. Nybergh, J. Himberg, K. Hjelt, "Contact Detection System for Mobile Terminals." In: International Symposium on Mobile Human Computer Interaction, Glasgow, Scotland, Sept 2004.

[P2] A. Ylisaukko-oja, E. Vildjiounaite, J. Mäntyjärvi, “Five-point Acceleration Sensing Wireless Body Area Network – Design and Practical Experiences.” In: IEEE International Symposium on Wearable Computers, Oct 2004.

[P3] J. Mäntyjärvi, S. Kallio, P. Korpipää, J. Kela, J. Plomp, “Gesture interaction for small handheld devices to support multimedia applications.” Journal of Mobile Multimedia, Rinton Press, Vol. 1(2), 2005, pp. 92–112.

[P4] J. Kela, P. Korpipää, J. Mäntyjärvi, S. Kallio, G. Savino, L. Jozzo, S. Di Marca, “Accelerometer based gesture control for a design environment.” In: Personal and Ubiquitous Computing, Springer, DOI 10.1007/s00779-005-0033-8, (online), 2005.

[P5] J. Häkkilä, J. Mäntyjärvi, “User experiences on Combining Location Sensitive Mobile Phone Applications and Multimedia Messaging.” 3th International Conference on Mobile and Ubiquitous Multimedia MUM2004. College Park, Maryland, USA, October 27–29, 2004, ACM Press.

[P6] G. Iacucci, J. Kela, P. Pehkonen, “Computational support to record and re-experience visits. Personal and Ubiquitous Computing.” Vol. 8 No 2, Springer-Verlag, 2004, pp. 100–109.

[P7] G. Iacucci, J. Kela, J. Plomp, “Configuring Gestures as Expressive Interactions to Navigate Multimedia Recordings from Visits on Multiple Projections.” 3th International Conference on Mobile and Ubiquitous Multimedia MUM2004. College Park, Maryland, USA, October 27–29, 2004, ACM Press. Pp. 157–164.

[P8] J. Mäntyjärvi, S. Kallio, J. Kela, P. Korpipää, “Enabling fast and effortless customisation in accelerometer based gesture interaction.” International Conference on Mobile and Ubiquitous Multimedia MUM2004. College Park, Maryland, USA, ACM Press. Pp. 25–31.

[P9] J. Häkkilä, J. Mäntyjärvi, "Collaboration in Context-Aware Mobile Phone Applications." HICSS 2005.

[P10] H. Strömberg, V. Pirttilä, V. Ikonen, "Interactive scenarios – building ubiquitous computing concepts in the spirit of participatory design." Personal and Ubiquitous Computing Vol. 8, No: 3–4, 2004, pp. 200–207.

[P11] H. Strömberg, J. Leikas, R. Suomela, V. Ikonen, J. Heinilä, (2005), "Multiplayer gaming with mobile phones – enhancing user experience with a public display." Intetain (Intelligent technologies for interactive entertainment); Madonna di Cambiglio 30.11.–2.12.2005, Italy.

[P12] J. Leikas, H. Strömberg, (2005), "Novel User Interfaces for Promoting Social Interaction in Multiplayer Games." ENACTIVE/05, November 17–18, Genoa, Italy.

[P13] J. Leikas, H. Strömberg, V. Ikonen, R. Suomela, J. Heinilä, (2006), "Multi-user Mobile Applications and a Public Display: Novel Ways for Social Interaction." Pervasive Computing and Communications, Pisa, Italy, 13–17 March 2006.

[P14] V. Ikonen, J. Leikas, H. Strömberg, (2005), "Nomadic Elderly: design requirements for technologically enhanced traveling experience." 5th International Conference of the International Society for Gerontechnology, Nagoya, Japan, May 24–27, 2005.

[P15] J. Mäntyjärvi, B. Gfeller, "Social Cliques: Group Awareness for Mobile Terminals." International Conference on Enactive Interfaces, Nov 2005, Genoa, Italy.

[P16] S. Kallio, J. Kela, J. Mäntyjärvi, J. Plomp, "Visualization of Hand Gestures for Pervasive Computing Environments." In: ACM Advanced Visual Interfaces, May 2006, Venice, Italy.

[P17] P. Korpipää, "Blackboard-based software framework and tool for mobile device context awareness." VTT Publications 579, Espoo, 2005. 225 p. ISBN 951-38-6669-6; 951-38-6670-X. http://virtual.vtt.fi/inf/pdf/publications/2005/P579.pdf.

[P18] J. Mäntyjärvi, "Sensor-Based Context Recognition for Mobile Applications." VTT Publications 511, Espoo, 2003. Ph.D. Dissertation. 118 p. + app. 60 p. ISBN 951-38-6253-4; 951-38-6254-2. http://virtual.vtt.fi/inf/pdf/publications/2003/P511.pdf.

[P19] J. Mäntyjärvi, H. Keränen, T. Rantakokko, "Sharing multimedia and context information between mobile terminals." In: Proceedings of the Workshop on Memory and Sharing of Experiences, In International Conference on Pervasive Computing, 2004.

[P20] J. Mäntyjärvi, J. Himberg, P. Kangas, "Sensor signal database and analysis for exploring context recognition of mobile devices." In: Proceedings of the Workshop Towards Benchmarks and a Database for Context Recognition. In: International Conference on Pervasive Computing, 2004.

22. Automation at Your Service in Future Homes (TUPA)

Veijo Lappalainen, Arto Laikari, Pekka Ala-Siuru, Kalevi Piira,
Hanna Strömberg and Timo Urhemaa

22.1 Project Goals

The starting point of the TUPA project was to define and implement items to the exhibition *Easy Life – Automation at Your Service* at the Finnish Science Centre Heureka, which was going to start in March 2004 for duration of one year. The exhibition was a part of Finnish Automation Society's 50 year anniversary.

The theme of the exhibition was extensively innovative and smart automation solutions from various fields starting from the industry to future homes. The general objective of the exhibition was to make automation and technology familiar to adults as well as to children regardless of their gender and especially inspire the growing generations to get interested in technology.

The goal of the project was to develop an actual integrated smart and intuitive future home environment and conduct a large user research with it in the Finnish Science Centre Heureka. It included the development of intuitive and interactive user interfaces as well as key applications and services for home automation. The following exhibition item ensembles were selected for implementation

- new ways to control home equipments as well as the adjustable table
- bus based home control system as well as future atmospheric lighting
- remote home management system and "is the dog in the house" (Radio Frequency Identification, RFID).

The idea was that with the aid of demonstrations, visitors would see the possibilities and principles of technology and essence of automation in a familiar environment – home. The targeted impact was also to improve knowledge of home automation market and business and to increase the knowledge and interest of home automation and home communication technologies among the masses. An additional targeted impact was to increase VTT's positive media visibility, to make VTT capabilities better known and to raise the general interest in VTT's activities, i.e. marketing VTT to the general public.

The itemized research targets of the project were to

- develop intuitive and interactive user interfaces for home automation.
- develop ICT-infrastructure and platform technologies.
- develop key applications and services for home automation.
- research the usability of home automation application and technical solutions.
- develop interdisciplinary technology.
- improve knowledge of home automation market and business.
- promote international cooperation.
- evaluate applications and technical solutions by prototypes.
- increase the knowledge and interest of home automation and home communication technologies among the masses.

22.2 How the Project Was Done

The project was implemented in co-operation between three different VTT research institutes by the following project team: Veijo Lappalainen (project manager) and Kalevi Piira from VTT Building and Transport, Arto Laikari and Timo Urhemaa from VTT Information Technology and Pekka Ala-Siuru from VTT Electronics.

The project included the industrial co-operation with following companies: Abloy, Aptus, Electrolux, Elektrobit, Elisa, EMFiTech, Fläkt Woods, Karin Metalli, Meritie, MohitoNet, Nokia, Ouman, Produal ja Smarthouse International.

Regarding funding, the TUPA project was divided into three parts: the preliminary study & roadmapping phase "TupaRoad", the research and development part of the implementation stage “TUPA VTT” and the industrial co-operation part of the implementation stage “TUPA Elisa”.

TupaRoad project was prepared, implemented and reported during the year 2003. The project was financed by Tekes, industry and VTT. The project developed a concept for home automation system architectures, key applications and user interfaces based on companies’ existing products, literature research and scenario analysis. A technology roadmap was also created. The project was also a prestudy to define exhibition items for Heureka’s automation exhibition, which were refined and further developed in the continuation sub-projects TUPA VTT and TUPA Elisa.

TUPA VTT and TUPA Elisa were implemented as one combined TUPA project under the same project plan. TUPA VTT was funded by VTT from the theme Future Communications Technologies and TUPA Elisa by co-operative industrial partners led by Elisa. The TUPA project started at the end of 2003 and it was completed in 2004.

Based on a Tekes funding proposal, the application for an actual research phase of TUPA project was changed into an industry-lead application, in which Elisa Oyj took the responsibility for the project leadership.

Although the project was creating novel home automation prototypes, demanding requirements existed for the implementation of exhibition items compared to standard research type prototypes. Exhibition items were going to be used by hundreds of thousands of visitors and they should be functioning several years without operators. They should also survive transportation to different countries. Exhibition ensembles should also be attractive, user friendly and easy to use regardless of visitors cultural backgrounds or language. Robustness and fault tolerance were also very important factors in implementation, because it was known that some visitors could be trying to misuse the items.

Some of the home automation concepts produced for home had to be also modified for the exhibition environment taking into account various requirements of a large general public event.

Multidisciplinary team of technology experts, visualists, carpenters and several other specialists were collected to define and implement the exhibition ensembles to ensure the quality of items.

22.3 Description of Results

The results of the project are classified as follows: the preliminary study and roadmapping, smart home environment demonstrators at Heureka and the usability studies based on the demonstrators.

The preliminary study and roadmapping

The outcome of the preliminary study and roadmapping phase was the research report titled *Future Home Automation at your Service; state of the art, scenarios and roadmap.* The report included the state-of-the-art study of industrial products, an analysis of advanced solutions for home automation, a scenario of a future home, a concept for future home automation, a market analysis and a technology road map for home automation.

Smart home environment demonstrator at Heureka

The specification and implementation of smart home environment's demonstrators for Heureka's *Easy Life – Automation at Your Service* exhibition included the following exhibition item ensembles E51–E55 (Heureka codes) made in co-operation with industrial partners:

- Mobile key (E51)
- Adjustable mood lighting (E52)
- Right amount of light (E53)
- Control your home (E54)
- Gesture control (E55).

The more detailed descriptions of the exhibition items follows later on.

Usability study using Heureka's demonstrators

The usability evaluation was carried out to collect user feedback concerning the exhibition stand "E54 Control your home". This prototype of a future home was implemented in the TUPA-project. The exhibition stand demonstrated the possibilities of technology and automation in future homes and peoples' daily lives.

End user involvement enabled evaluating the usefulness and acceptability of the implemented home appliances as well as the seamlessness of the interaction with the functionalities. The users evaluated the individual interfaces and the integrated home control system where the appliances could be monitored and controlled. In addition, users' opinions concerning future living in general were collected.

10 test users validated the home appliances independently. The participants were asked to perform small tasks and answer to predefined questions about the appliances. Additionally, user observation in the exhibition stand was carried out. The evaluation results and user response are presented in the evaluation report.

As a summary, the feedback of the user action should be visible and easy to understand. This concerns both the appliance interfaces and the integrated home control system. The users had difficulties using the functionalities that require complicate and multi-phased actions. The functionalities that bring concrete benefit were seen as the most useful. The advantages of remote observation of the home were emphasized, e.g. the "Burglar alarm" function was found useful. "Is the dog in the house?" was one of the appliances that was found nice and appealing in the exhibition stand. However, its usefulness in the real life was not rated high since the users could not foresee themselves using this kind of appliance in practise.

According to the test users, the home control system and its functionalities should be sustainable and reliable. In general, the results show that the users want that their homes can be controlled but they want to define the controllable objects themselves. Correspondingly, the users like the possibility of controlling their homes but they do not want to be monitored themselves. The idea of being monitored at home (e.g. "Camera surveillance") raised several ethical questions among the users.

Descriptions of the exhibition items ensembles

E51 MOBILE KEY

The mobile key demonstrates how user identification can be linked to be a function of an electric key. In comparison with traditional keys, the user identification function makes the electric key a much more secure solution.

A mobile phone can be used as an electric key for locks that have multiple users. Electric keys would be useful, for example, in telecom facilities and perhaps in office premises. Electric keys are easier to control than mechanical keys. If necessary, the key also enables identification of the user. This ensures that someone finding the key would be unable to use it, because he or she would not know, for example, the rhythm required to open the lock. Figure 1 illustrates the demonstrator.

Figure 1. Mobile key (E51).

An electric key would also be useful in other places that require user identification, such as logging onto a computer.

This demonstrator was produced in mutual co-operation between Elisa, Abloy and VTT.

E52 ADJUSTABLE MOOD LIGHTING

Every source of light has a spectrum, which determines what colour the light will be, and how it reproduces colours of various objects. Daylight is bluer than the yellowish light of an incandescent bulb. However, in suitable circumstances, we do not experience the shade of the existing light as having colour, but as being neutral.

The shade of colour can be adjusted in this new kind of light fixture. It has different buttons for adjusting partial colours such as red, green, and blue. The combination produced by lamps may remind the observer of the yellow of the sun, ice blue, the glow of sunset, or perhaps even the green of summer trees.

The demonstrator is initiated by pushing the start button. With the adjustment levers, the lighting in the lounge area can be adjusted to suit the visitor's preference. The chosen colour value is shown on the computer's colour illustration.

Figure 2. Adjustable mood lighting (E52).

The choice can be saved and compared to the choices of other visitors.

This demonstrator was produced in co-operation between VTT and Idman Oy.

E53 RIGHT AMOUNT OF LIGHT

In the near future it will be commonplace to control lighting with doorway switches, wireless remote controls or mobile phones. You can turn a light on by sitting in your armchair, which adds to comfort and makes life easier for the elderly and people with special needs. Controlling appliances can be automated. Lights and air conditioning can be turned on with the aid of a movement detector. This can also be done through the Internet with a computer, mobile telephone, and other mobile data terminal equipment. In scene lighting, large units that are composed of smaller devices are controlled with one button. New technology makes the use of lighting easier and reduces unnecessary energy consumption. Figure 3 illustrates the demonstrator.

Figure 3. Right amount of light (E53).

This demonstrator was produced in co-operation between VTT and Schneider Electric Oy.

E54 CONTROL YOUR HOME

Exhibition item ensemble E54 was constructed from two connected and interoperating sub-emsembles: "*Is the dog in the house*" and "*Control your home*".

Is the dog in the house?

Three stuffed toys are equipped with RFID (Radio Frequency Identification) tags which can be used to follow the toys' movements with RFID readers located in certain places. The readers in this exhibit are located in doghouses, feeding areas, and dog-owners' homes. Each dog has its own place, which you must find by "trial and error". A light will turn on and dog will bark happily, when you have found the right "home locations" for each dog. If you put a dog in the wrong location (e.g. home, dog house or feeding cup), dog will whine sadly revealing a negative selection. The control screen shows the location of the dogs as well. Figure 4 illustrates the demonstrator.

Figure 4. Is the dog in the house (E54a).

This items main themes are RFID technology, recognition and localization. The RFID tags inside the dogs and the readers hidden in structures illustrate how recognition and localization can be made automatically. The subsystem is also connected to home control panel, where dog locations can be viewed remotely on a home touch screen control monitor.

Control your home

Different home appliances can be joined into an integrated home control system where they can be monitored and controlled either from the home or from almost anywhere. Did you forget to turn off the stove? Is the front door locked? Where are the pets? These are questions to which one would be happy to have answers at any time. It is possible with a device connected to any Internet or mobile telephone network, granted one has a control system at home that supports it. The programmable control system allows one to monitor and control doors, windows, and all electrical appliances at home. By using a touch screen monitor, occupants may be alerted about various things such as smoke, moisture, and burglaries. Figure 5 illustrates the demonstrator.

Figure 5. Control your home (E54b).

This demonstrator includes the following functions: carport, burglar alarm, camera surveillance located near the motion detectors, camera surveillance at the front door, motion detector, humidity, fire alarms, temperatures, outer door control and monitoring, aquarium, windows, the stove, refrigerator/freezer, lights and a control-monitor touch screen.

The demonstrator was produced in co-operation between VTT and Smarthouse International Oy.

E55 GESTURE CONTROL

Controlling home appliances with just a simple movement of your hand is possible with so-called gesture connection technology, which is based on a wireless control device the size of a bar of hotel soap. The device in question is the SoapBox (Sensing Operating Activating Peripheral Box). Its sensors identify human gestures learned beforehand and in this way wirelessly control devices and objects in the environment. A person can control home appliances, such as television or Internet connection, with hand gestures from anywhere inside the home. Figure 6 illustrates the demonstrator.

Figure 6. Gesture control (E55).

This demonstrator was produced in co-operation between VTT, Meritie Oy and University of Lapland.

Co-operation and networking

The following companies were involved in the project during the preliminary study & roadmapping phase:

Abloy Oy, Aptus Sensible Electronics Ltd, Electrolux, Elektrobit Group Oyj, Elisa Oyj, EMFiTech Oy, Fläkt Woods Oy, Idman Oy, Karin Metalli Oy, Meritie Oy, Mohito Networks, Nokia Mobile Phones, Ouman Oy, Produal Oy, Schneider Electric Oy, Smarthouse International Oy.

The collaborating companies during the implementation phase of the demonstrators are named in the paragraph "How the project was done".

Easy Life – Automation at Your Service exhibition is unique for its theme and has received plenty of interest abroad. After being the first year in Finland at Heureka, the exhibition was hired to Lisbon, Portugal. After that the exhibition will go next to Ciudad de Mexico, Mexico. In this way the exhibition and its contributors will receive also international visibility.

22.4 Conclusions

The project managed to fulfil its requirements very well. The most important goal was to create a representing and extensive exhibition ensemble to Heureka's automation exhibition, which fulfils all demanding requirements set for the implementation. This goal was reached well. The project also created several new co-operation networks and strengthened VTT's home automation know how, which has later generated several new internal and external projects on that area. Other itemized results include:

- Co-operation and interaction with several companies working at intelligent home area has been created.
- The created networks enable the emergence of new industrial projects.
- There are several possibilities to create novel products and concepts used in demonstrators to be further developed in future projects.
- Participation in Heureka's "Easy life" exhibition will spread awareness about VTT and its work widely.
- VTT's media visibility has been increased.

Project implementation environment generated also good possibilities to conduct usability studies on home automation. According to the projects usability study, the home control system and its functionalities should be sustainable and reliable. In general, the results show that the users want that their homes can be controlled but they want to define the controllable objects themselves. Correspondingly, the users like the possibility of controlling their homes but they do not want to be monitored themselves. The idea of being monitored at home (e.g. "Camera surveillance") raised several ethical questions among the users.

Publication

[P1] P. Ala-Siuru, A. Laikari, V. Lappalainen, T. Urhemaa, “Future Home Automation at your Service – State-of-the-Art, Scenarios and Roadmap.” 2004. 71 p. (In Finnish) http://www.vtt.fi/inf/julkaisut/muut/2003/tuparoad_raportti.pdf.

Series title, number and report code of publication

VTT Publications 629
VTT-PUBS-629

Author(s)
Sipilä, Markku (ed.)

Title

Communications Technologies
VTT's Research Programme 2002–2006
Final Report

Abstract

The results of VTT's strategic technology theme (programme) "Future Communications Technologies" are described. The theme was a research entity during the years 2002–2006 with a goal to study and develop technologies which enable users to communicate "anywhere, anytime" with the appropriate quality of service. This goal was approached on several hierarchical levels from physical electronics up to software based systems. Therefore the theme consisted of three main topical areas: networks, radio frequency technology and smart environments. At radio frequencies new kinds of micro-electromechanical (MEMS) components and devices were developed, as well as selected manufacturing processes and integrated circuits. In network technology the emphasis was on the new generation mobile network with wireless multimedia applications. Basic technologies for intelligent environments were also researched and developed.

ISBN
978-951-38-6997-7 (soft back ed.)
978-951-38-6998-4 (URL: http://www.vtt.fi/publications/index.jsp)

Series title and ISSN
VTT Publications
1235-0621 (soft back ed.)
1455-0849 (URL: http://www.vtt.fi/publications/index.jsp)

Project number

Date
March 2007

Language
English, Finnish abstr.

Pages
354 p.

Name of project

Commissioned by

Keywords
telecommunication, communication technologies, wireless communication, mobile networks, autonomic networks, radio networks, cellular networks, service architecture, filter technologies, nomadic media, future homes

Publisher
VTT
P.O.Box 1000, FI-02044 VTT, Finland
Phone internat. +358 20 722 4404
Fax +358 20 722 4374

Julkaisun sarja, numero ja raporttikoodi

VTT Publications 629
VTT-PUBS-629

Tekijä(t)
Sipilä, Markku (toim.)

Nimeke

Tiedonsiirtoteknologiat
VTT:n teknologiateema 2002–2006
Loppuraportti

Tiivistelmä

VTT:n strategisen teknologiateeman (tutkimusohjelman) "Tulevaisuuden tiedonsiirtoteknologiat" tulokset on kuvattu tässä raportissa. Teema oli vuosina 2002–2006 toteutettu tutkimuskokonaisuus. Sen päämääränä oli tutkia ja kehittää teknologioita, jotka mahdollistavat käyttäjälle riittävän laadukkaan tiedonsiirron "missä ja milloin tahansa". Päämäärää lähestyttiin useilla hierarkiatasoilla fysikaalisesta elektroniikasta ohjelmistokeskeisiin järjestelmiin. Tämän mukaisesti teema koostui kolmesta pääaihealueesta: verkot, radiotaajuusteknologia ja älykkäät ympäristöt. Radiotaajuisen teknologian alalla kehitettiin uudenlaisia mikrosähkömekaanisia (MEMS) komponentteja sekä myös valmistusprosesseja ja integroituja piirejä. Verkkoteknologiassa pääpaino oli uuden sukupolven matkaviestinverkossa ja langattomissa multimediasovelluksissa. Tutkittiin ja kehitettiin myös perusteknologioita älykkäiden ympäristöjen tarpeisiin.

ISBN
978-951-38-6997-7 (nid.)
978-951-38-6998-4 (URL: http://www.vtt.fi/publications/index.jsp)

Avainnimeke ja ISSN	Projektinumero
VTT Publications 1235-0621 (nid.) 1455-0849 (URL: http://www.vtt.fi/publications/index.jsp)	

Julkaisuaika	Kieli	Sivuja
Maaliskuu 2007	englanti, suom. tiiv.	354 s.

Projektin nimi	Toimeksiantaja(t)

Avainsanat	Julkaisija
telecommunication, communication technologies, wireless communication, mobile networks, autonomic networks, radio networks, cellular networks, service architecture, filter technologies, nomadic media, future homes	VTT PL 1000, 02044 VTT Puh. 020 722 4404 Faksi 020 722 4374

VTT PUBLICATIONS 629 Communications technologies. VTT's Research Programme 2002–2006

A major communications research effort at VTT Technical Research Centre of Finland took place in the strategic technology theme (programme) "Future Communications Technologies". The results of that work within the timeframe 2002–2006 are described in this volume.

The goal was to study and develop technologies which enable users to communicate "anywhere, anytime" with the appropriate quality of service. This goal was approached on several hierarchical levels from physical electronics up to software based systems. The theme accordingly consisted of three main topical areas: Networks, Radio frequency technology and Smart environments. At radio frequencies new kinds of microelectro-mechanical (MEMS) components and devices were developed, as well as selected manufacturing processes and integrated circuits. In network technology the emphasis was on the new generation mobile network with wireless multimedia applications. Basic technologies for intelligent environments were also researched and developed.

Julkaisu on saatavana	Publikationen distribueras av	This publication is available from
VTT PL 1000 02044 VTT Puh. 020 722 4404 Faksi 020 722 4374	VTT PB 1000 02044 VTT Tel. 020 722 4404 Fax 020 722 4374	VTT P.O. Box 1000 FI-02044 VTT, Finland Phone internat. + 358 20 722 4404 Fax + 358 20 722 4374

ISBN 978-951-38-6997-7 (soft back ed.)
ISSN 1235-0621 (soft back ed.)

ISBN 978-951-38-6998-4 (URL: http://www.vtt.fi/publications/index.jsp)
ISSN 1455-0849 (URL: http://www.vtt.fi/publications/index.jsp)