JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

This special issue of the *Journal of Telecommunications and Information Technology* contains selected papers from the conference organized in the frames of the Sixth Framework Program Project *Academic Internet Television Network Showcases: the Best of Good Practice Activities* – ATVN-EU-GP (http://www.atvn-eu-gp.pl). A consortium of eight partners: the National Institute of Telecommunications (Poland), the ICM, Warsaw University (Poland), Sofia University (Bulgaria), the Institute of Fundamental Technological Research, Polish Academy of Sciences (Poland), the Centre for Administration and Operations, the Academy of Sciences of the Czech Republic (Czech Republic), Starptautiska Lietiskas Optikas Biedriba (Latvia), the Technical University of Kosice (Slovakia) and Lviv Centre for Scientific, Technical and Economic Information (Ukraine), initiated the implementation of this project. The main goal of the project was to stimulate, encourage and facilitate the participation of teams from the new member states (NMS), the associated candidate countries (ACC) and the newly independent states (NIS) in research projects and activities in the areas of information society technologies (IST), such as: e-business, e-government, e-learning and e-health. The project aimed at identifying, exploiting and further improving the good IST practices in order to use them as showcases for strengthening the IST RTD activities in these countries. Its strategic objectives were also to stimulate, encourage and facilitate the participation of organization from these countries in future IST research and good practice projects and to increase their innovations and competitiveness in the European business and industry. The main focus was put on promotion of research competencies and establishment and reinforcement of networks of research organizations from the NMS and the ACC with organizations from the other member states. The ATVN-EU-GP conference was among the main instruments for achieving this. The conference, held in Pułtusk, Poland, in December 2005, gave the unique opportunity for presentation of the twelve selected in the first stage of open contest best practice projects within the four mentioned IST areas. During the conference the four best projects have been selected in the second stage of the contest (one in each selected area). The best of ”Good Practices Projects” were prize-winning by production of audio/video programs, which will be and placed on the new multimedia on-demand library as a beginning of the archive (http://www.atvn.pl/atvn-eu-gp/index1.php) showcases the best of good practices projects from EU.

This issue, titled *European IST research enlargement*, includes a selection of nine interesting and valuable papers, presented at the conference. All selected papers have their origin from some EC RTD projects and international co-operation and networking. Although the call for papers was in the four mentioned above IST areas, most of the contributions are presenting research and development outcomes that cross the thin boundaries between these areas.

The e-business section contains two papers. The first one, *MUMMY – mobile knowledge management*, presents the achievements of the MUMMY project which aimed to research and develop means to improve the efficiency of mobile business processes through mobile and personalized knowledge management. The authors come from Department of Computer Science and Engineering, Czech Technical University in Prague and ZGDV – Computer Graphics Center, Darmstadt. MUMMY aimed to research and develop means to improve the efficiency of mobile business processes through mobile, personalized knowledge management. The project resulted in a set of integrated components forming the MUMMY system. The system can be used by companies for providing telemedicine and e-learning services as well.

The second paper in this section, *Empowering the mobile worker by wearable computing – wearIT@work*, deals mostly with the IST priority area e-work. The paper presents some early achievements of the wearIT@work project, the largest project world-wide in wearable computing.

The intention of wearIT@work is to explore some applications of computer systems embeded in clothes in various industrial environments. The research is based on the user-centred design approach and scenario definitions which are to be validated. The project aims to deliver a new technology having a long-term impact on the organization of individual and collaborative work and creating new ways to organize work.

The e-government section contains four papers. The first one, *EWD-P as an example of a "The Best of Good Practice" project*, presents the features of an electronic document exchange systems developed by the Polish company Rodan Systems SA. It could facilitate the decision-making process of the Polish government related to a great number of legislative issues and cases on an EU level. The EWD-P system integrates intelligent workflow management functionality aiming at supporting a complex flow of documents at the Polish state administration.

The paper *Innovative approach to identity management solution development for e-government at EU level* presents some achievements of the Sixth Framework RTD Integrated Project GUIDE. The project consortium includes partners from Bulgaria, Estonia and Hungary. The project aims at developing an open identity management architecture for e-government solutions in Europe. The paper presents the interdisciplinary perspective of the problem, bringing together internationally recognized expertise in technology, institutional, policy, socio-economic, legal and ethical research. It analyses the specific needs of the EU based upon the social, ethical and legislative differences regarding privacy and data protection. The project seeks to overcome the fragmentation of existing identity management initiatives.

The next two papers in this section deal with research and development projects and initiatives related to integration of people with special needs into the information society. This papers cover a variety of research and development areas, including e-government, e-learning, e-inclusion, e-health, etc. The paper *A way of integrating deaf, hearing- and speech-impaired people into modern communication society* puts in its focus the concept of the Relay Center Austria, which is a service and competence centre, for deaf, hearing- and speech-impaired people. The paper considers the issues both from organizational and business perspective.

The paper *Inclusion of sign language users via information and communication technology* analyses the differences between the sign and spoken languages and some adequate technical solutions appropriate for wider use of sign languages. Some intellectual, scientific, organizational, and political barriers for the necessary developments are discussed as well.

The e-health section contains two papers. The first one *Implementation of information security management system in the small healthcare organization* presents the outcomes of a subproject developed in the frames of the DIGI-Q course, part of the project "Quality and On-line Confidence in SMEs e-Business Processes" funded by the EC. The paper deals with topics of methodology for data protection, audit and certification, system, risk assessment and risk management, and their implementation in small healthcare organizations.

The next *A roadmap towards healthcare information systems interoperability in Greece* presents results of a comprehensive assessment, concerning the situation in the area in e-health in Greece and defined recommendations for a successful implementation of a national e-health system.

The e-learning section contains only one paper *Comparative analysis of models and platforms for the e-learning portals*. Author analyses the existing models and platforms for e-learning in terms of performance, user oriented services, data management, portal adaptors and web-infrastructure, etc. Some useful guidelines are also available, revealing the advantages and disadvantages of compared systems and giving recommendations for customers in making a choice among these platforms.

The guest editor would like to thank here all the authors for their contributions and the reviewers for their hard work in preparing their submissions, reviewing, and revising the papers on time.

Roumen Nikolov
Guest Editor

Paper

MUMMY – mobile knowledge management

Martin Klima, Zdenek Mikovec, Pavel Slavik, and Dirk Balfanz

Abstract— The project MUMMY funded by the European Commission develops means to improve the efficiency of mobile business processes through mobile, personalized knowledge management. MUMMY approaches the challenges of modern mobile work processes. To do so, it takes advantages of latest achievements in mobile connectivity and its capabilities (like "always on-line" high bandwidth personalization ubiquity), latest hardware options like camera-equipped handheld devices, and uses multimedia, hypermedia, and semantic web technologies. Technical development and appliance of the results are intensively consulted and integrated with business processes of several commercial organizations that are members of the MUMMY consortium. In this paper the achievements of MUMMY are introduced and individual components are briefly described.

Keywords— knowledge management, mobile computing, collaboration, multimodal interaction, SVG, RDF, context of use, facility management.

1. Introduction

A large, and still rising percentage of workers is doing some mobile business work. From the IT point of view, one of the immediate consequences is the increasing demand for being integrated into the global and corporate network both in case of recording new information and retrieving needed information.
The MUMMY[1] was a project, funded by the European Commission (EC) and the Federal Office for Education and Science (BBW) in Switzerland, to research and develop means to improve the efficiency of mobile business processes through mobile, personalized knowledge management.

The project resulted in a set of integrated components forming the MUMMY system. This enables mobile workers to never loose track and coherence of their mobile knowledge and provides:

- New approaches to mobile information and knowledge handling by extending knowledge management (KM) to support authoring, sharing, retrieval, and visualization processes in mobile work.
- Added value by just-in-time mobile assistance, facilities to speed up the workflow of spatially distributed business processes.
- Savings through efficient/natural forms of on-the-spot information capturing.
- Avoidance of delayed data entry and post-edit.
- Support for mobile cooperative (planning) processes.

[1]See, www.mummy-project.org

The activities in the project are concentrated on:

- New approaches to mobile information and knowledge handling. This will be achieved by extending KM to support key processes in mobile work.
- Just-in-time mobile assistance supported by context aware information processing in collaborative environment.
- Real-world experience and background through trials and work in the application areas of facility management, the building trade and the service domain.

The following chapters describe the functionalities of the individual components of the MUMMY system that implement the named functionalities.

2. Main features

The MUMMY system is a generic platform to support knowledge acquisition and retrieval in mobile business processes. The system and its components are able to support various areas of use. MUMMY focuses on mobile support in the areas of facility management and technical services. In particular the MUMMY core application (a specific adaptation of the MUMMY platform), specified in cooperation with ARCADIS GmbH, addresses the needs of mobile facility management.
In general MUMMY enables mobile workers to never loose track and coherence of their mobile knowledge, which means the MUMMY system meets the needs of mobile workers to **capture and access knowledge coherently** by an information-in-context approach using multimedia annotation and location/task-related assistance.
The main features of the system can be shortly summarized into the following categories:

- contextual behavior,
- data adaptation,
- advanced information searching,
- scalable vector graphics (SVG) collaboration, real time SVG annotation sharing,
- hypervideo,
- Grenoble components,
- problem tracking.

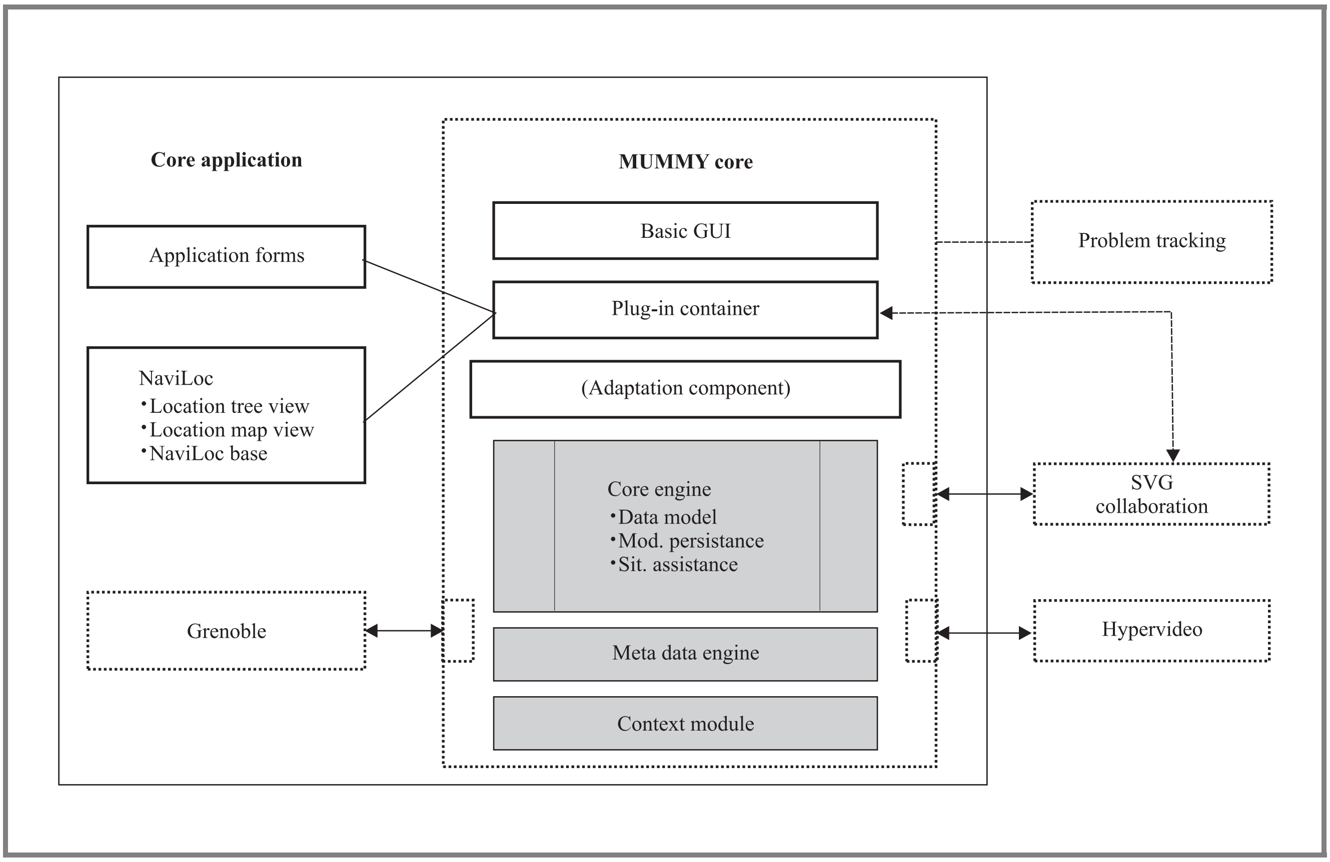


Fig. 1. MUMMY system components.

The named features were implemented by individual partners in the MUMMY project and finally integrated into one system (Fig. 1)

2.1. Contextual behavior

Nowadays most mobile workers document their insights in terms of paper-based notes and sketches and enter them into their databases when they are back in the office. MUMMY utilizes the new multimedia facilities in personal digital assistants (PDAs) to record the worker's comments. These comments in a form of photos, hand-drawn sketches, speech notes and simplified input forms are added to existing multimedia documents (such as videos with predefined hotspots, SVG drawings, maps and site plans). Thereby, the MUMMY system recognizes and **automatically correlates new recordings with descriptive information** from user's context (such as time, locations, people, tasks, and projects) in order to bring structure into the unstructured files and prepare them for advanced (semantic) search.

2.2. Data adaptation

User's context is utilized for a range of functionalities. Data adaptation is one of the representative use cases. The data adaptation process adjusts the application data to the current context of use utilizing information about environment, device, user and additional meta-information describing the multimedia data itself. The adaptation system is responsible for adapting the requested information to the current context of use and specified user query. For more efficient adaptation, besides the current context, the adaptation system utilizes additional semantic description of the adapted data (Fig. 2).

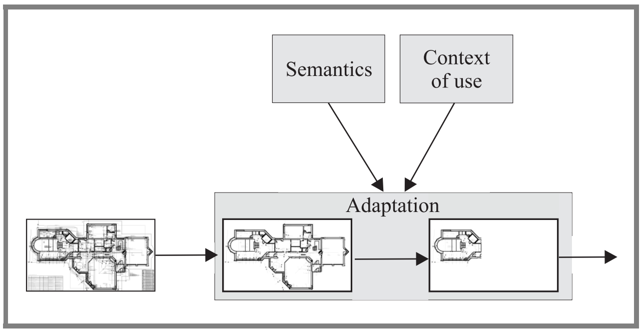


Fig. 2. Data adaptation process.

The result of the adaptation is data that fit the requirements of the user and the given context of use. The adaptation reduces the amount and complexity of delivered

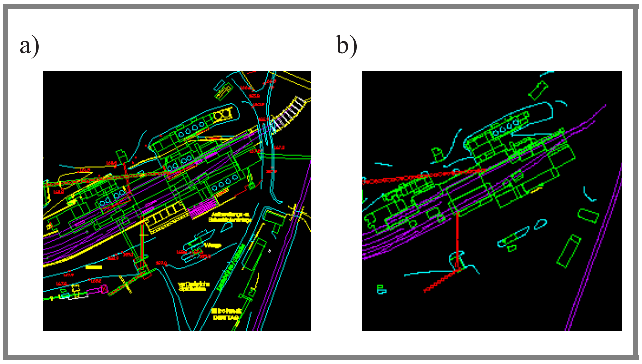


Fig. 3. Adaptation result example: (a) original document 2 439 objects, 376 kB; (b) filtered document 726 objects, 173 kB.

information so that it is usable in the mobile environment (Fig. 3).

2.3. Advanced information searching

During the data retrieval process in both the mobile and the stationary environment, the MUMMY system can utilize the semantic information for building a semantic network (metadata) for faster data retrieval. Users of the system have then the possibility to request data without precisely knowing their location. The querying can be also based on relation to other data, event or location. Building of the semantic network is done automatically as soon as the user creates the data, for example audio recording, photo, textual annotation, meeting schedule and so on (see Fig. 4 for

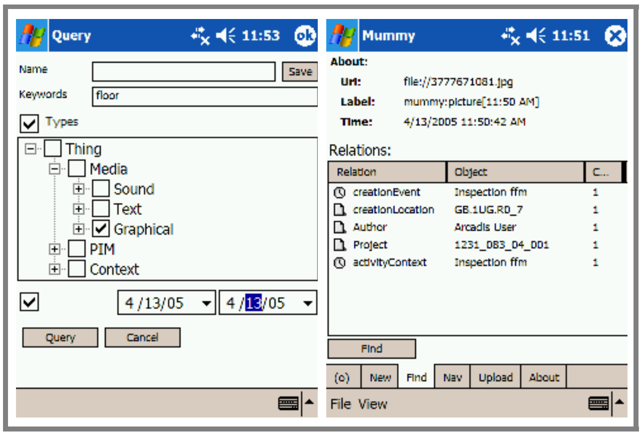


Fig. 4. Semantic data query tool.

semantic data query tool). The semantic information (metadata) is stored in a resource description framework (RDF) based database.

2.4. SVG collaboration, real time SVG annotation sharing

The goal of task collaboration and live sharing of annotations is to provide a tool and corresponding infrastructure for easier decision making and storing of acquired knowledge in mobile environment by enabling collaborative work. The system provides on-line data access to corporate database while using a variety of mobile devices like PDAs, notebooks and smart phones. As a basic data format for browsing construction plans in the mobile environment, SVG data format was chosen. This data format is designed for presenting of vector graphics and other hypermedia content on the web. The MUMMY architecture consists of mobile SVG editor (see Fig. 5), client network

Fig. 5. Mobile SVG editor.

component based on Jabber protocol and a server side collaboration and data sharing component. The collaborative architecture is depicted in Fig. 6.

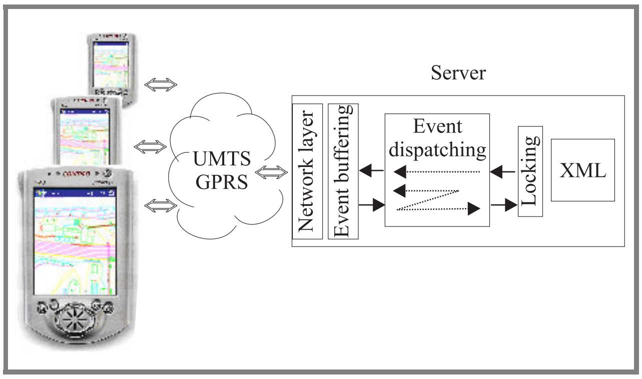


Fig. 6. Collaborative SVG system architecture.

The collaborative component is especially designed for the mobile environment and is therefore able to deal with the mobile network unreliability. Users can continue in work during short time connection failure. The changes made in the SVG plans are then automatically synchronized and distributed to all connected clients.

2.5. Hypervideo

Like SVG collaboration, the hypervideo module can be used either as a stand-alone application, or as an integrated part together with the MUMMY core and the MUMMY core application. Hypervideo provides a complete set of functions to **use and author interactive video sequences**. Thereby, hypervideo combines details in a video sequence with other documents, which can be any form of multimedia (images, audio, video, text). These links to other media appear as visible and "clickable" regions within the authored video. Video annotation may emphasise specific parts of the video itself or the video material along a story line or a spatial path (e.g., the path of a site inspection). The latter case is integrated with the MUMMY core by interfacing the time/task/location/user-related management of inter-related information parts in context.

Finally, the hypervideo users can annotate video content during a presentation by combining certain details in the video sequences with own additional information in order to share their knowledge with others in the community. This collaborative use of hypervideo can be considered as a dynamic document, which can be modified and extended by conversations and discussions regarding group activities.

2.6. Grenoble

Grenoble supports in particular **pre- and postprocessing** of information needed or collected on-site during, e.g., a site inspection.

For preprocessing purposes it facilitates the interrelation of locations with tasks and media. These tasks and media will then be made available pro-actively on-site under the defined activation conditions, such as user's arrival at a certain spot.

In the postprocessing phase, visualization of interrelations is provided for both previously authored and mobile generated info items. The user is enabled to review consistency and completeness of "the information in context" and may alter and correct relations and data elements. Furthermore Grenoble provides a **report-generator**, which is able to export mobile collected inspection results XML-based into the customers report templates (e.g., MS Word Templates).

The main visualization methods and approaches to manipulate the interrelated data-space are not domain specific. In this regard Grenoble may be used as an external module extending the usability of the MUMMY core. However, as Grenoble is in particular designed for interaction with the MUMMY data models it can not be used completely stand-alone like the SVG and hypervideo modules.

2.7. Problem tracking

The problem tracking system facilitates the day-to-day operational tasks of field technicians in the technical service (TS) domain. It implements the current real-life technical service procedures employed in INTRACOM's TS department. Currently it operates as a standalone application but it can be integrated with the MUMMY core and use MUMMY core functionality. Main functionalities of this application are:

- Transparent information exchange between the TS "headquarters" and the field technicians based on bi-directional synchronization of the mobile and home data repositories. Synchronization can be performed *on-user-demand* and is optimized for reduced bit-rate data services (like GSM/GPRS).
- Representation of context information {*user* + *time* + *task* (*product serviced*)/*location* (*where available*)} within its local database. This information can be used for automatic creation of context awareness and situation assistance information from the MUMMY core without any need for preprocessing.

3. The MUMMY use

As pointed out above, the MUMMY system is a generic platform to support knowledge acquisition and retrieval in mobile business processes. Potential MUMMY applications may therefore serve in a multitude of different business areas as a mobile knowledge assistant. In this section some possible ways to use MUMMY and its components shall be described. The main use of MUMMY is the mobile support for facility management.

Video-based e-learning. The hypervideo component can be used to support collaborative e-learning. A group of persons can interact remotely with interactive videos. These contain besides their obvious content clickable regions, which refer to more detailed information. The learning group can add further elements in the process of learning, can attach as well public visible questions and can track the history of discussion. All added hyperlinked information elements and discussion contributions can be searched and the referring part within the video can be jumped to immediately.

Mobile health-care support. MUMMY system can be used by companies for the provisioning of telemedicine service in order to provide remote medical help from nursing personnel and doctors to patients in distant locations. SVG collaboration can provide a very effective and intuitive way of working on X-rays images or other medical examinations (MRI, arteriography, ultrasound). The doctor in the headquarters will not only be informed at once but he/she can interact online on the image and may ask for more info from the local physician. All the files concerning the patient will not be lost or distributed across different hospitals but will be gathered in a central database consisting a very detailed and thorough medical history.

4. Conclusion

The MUMMY system has been implemented and individual components as well as the integrated system tested in

facility management scenarios. Currently the system is implemented as a functional prototype and members of the MUMMY consortium are considering application of the system in commercial sphere. Further development is expected in the follow-up projects.
The scientific results of the project were published in a number of conferences and scientific magazines (see References).

5. Project partners

The project MUMMY partners are:

- Zentrum für Graphische Datenverarbeitung e.V.
  Fraunhoferstr. 5
  64283 Darmstadt, Germany
  http://www.zgdv.de
- INTRACOM SA
  Markopoulou Ave.
  19002 Peania, Greece
  http://www.intracom.gr
- University of Applied Sciences
  Waedenswil
  Grüental, Postfach 335
  8820 Wädenswil, Switzerland
  http://www.hsw.ch
- Czech Technical University in Prague
  Computer Graphics Group
  Karlovo namesti 13
  12135 Praha 2, Czech Republic
  http://www.cgg.cvut.cz
- COSMOTE
  44 Kifissias Ave.
  15125 Athens, Greece
  http://www.cosmote.gr
- ARCADIS
  Europaplatz 3
  64293 Darmstadt, Germany
  http://www.arcadis.de

References

[1] J. Jelinek, "Derivation of information from structural and semantic description", Master thesis, Czech Technical University in Prague, 2003.

[2] Z. Mikovec, "Formal description of human-computer graphical communication", Postgraduate Study Report DC-PSR-2002-06, Czech Technical University in Prague, 2003.

[3] M. Grimm, M. R. Tazari, and D. Balfanz, "Towards a framework for mobile knowledge management", in *Proc. Pract. Asp. Knowl. Manag. PAKM*, Vienna, Austria, 2002, pp. 326–338.

[4] M. Finke, "Interactive video supporting CSCL environments", in *World Conf. E-Learn. Corp., Govern., Healthc., High., Educ. E-Learn 2002*, Montreal, Canada, 2002.

[5] D. Balfanz and J. Schirmer, "MUMMY: mobile knowledge management", *Comput. Graphik Top.*, issue 4, 2002.

[6] M. Finke, M. Grimm, and M. R. Tazari, "Design principles for a collaborative hypervideo user interface concept in mobile environments", in *Proc. 10th Int. Conf. Hum. Comput. Interact.*, Crete, Greece, 2003.

[7] M. Grimm, M. R. Tazari, and M. Finke, "User interface techniques for mobile agents", in *Proc. 10th Int. Conf. Hum. Comput. Interact.*, Crete, Greece, 2003.

[8] M. R. Tazari, "A context-oriented RDF database", in *Proc. First Worksh. Sem. Web Databas.*, Berlin, Germany, 2003.

[9] M. R. Tazari, M. Grimm, and M. Finke, "Modelling user context", in *Proc. 10th Int. Conf. Hum. Comput. Interact.*, Crete, Greece, 2003.

[10] C. Waldeck, D. Hess, and D. Balfanz, "Mobile liquid information spaces", in *Proc. Mensch Comput. "Interaktion in Bewegung"*, Stuttgart, Germany, 2003.

[11] Z. Mikovec and P. Slavik, "A structured approach to the interaction in 3D", in *Proc. 10th Int. Worksh. Hum. Comput. Interact.*, New Jeresey, USA, 2003, vol. 4, pp. 123–127.

[12] Z. Mikovec, M. Klima, and R. Foldyna, "GUI for graphical data retrieval by means of semantic filtering", in *Proc. 10th Int. Worksh. Hum. Comput. Interact.*, New Jeresey, USA, 2003, vol. 2, pp. 193–197.

[13] P. Zikovsky, Z. Mikovec, and P. Slavik, "A universal approach to multimodal user interfaces", in *Proc. 10th Int. Worksh. Hum. Comput. Interact.*, New Jeresey, USA, 2003, vol. 2, pp. 821–825.

[14] J. Jelinek, Z. Mikovec, and P. Slavik, "Declarative way of filtering of graphical data in mobile computing", in *Proc. 4th Int. Worksh. Mob. Comput. IMC 2003*, Stuttgart, Germany, 2003, pp. 38–45.

[15] R. Zenka and P. Slavik, "Panorama sketcing for PDAs", in *Proc. 4th Int. Worksh. Mob. Comput. IMC 2003*, Stuttgart, Germany, 2003, pp. 1–8.

[16] M. R. Tazari and K. Plöler, "User-centric service brokerage in a personal multi-agent environment", in *Int. Conf. Integr. Knowl. Intens. Multi-Agent Syst. IEEE KIMAS'03*, Cambridge, USA, 2003.

[17] C. Waldeck and D. Balfanz, "Mobile liquid 2D scatter space (ML2DSS) – a visual interactive information space (ispace) for displaying large amounts of information", in *8th Int. Conf. Inform. Visual.*, London, United Kingdom, 2004.

[18] M. Finke and D. Balfanz, "A reference architecture supporting hypervideo content for ITV and the Internet domain", *Comput. Graph.*, vol. 28, no. 2, pp. 179–191, 2004.

[19] T. Chambel, C. Zahn, and M. Finke, "Hypervideo design and support for contextualized learning", in *4th Int. Conf. Adv. Learn. Technol.*, Joensuu, Finland, 2004.

[20] Z. Mikovec, M. Klima, and P. Slavik, "Semantic driven visualization", in *Proc. CODATA Int. Worksh. Inform. Visual.*, Prague, Czech Republic, 2004.

[21] Z. Mikovec and P. Slavik, "Editing of 2D graphical information in mobile environment", in *Proc. Worksh.*, Prague, Czech Republic, 2004, pp. 198–199.

[22] D. Wiegand and D. Balfanz, "MUMMY – mobiles facilty management with mobile computers", in *Schweizer Energiefachbuch 200*, R. Köhler, Ed. Bern: 2004.

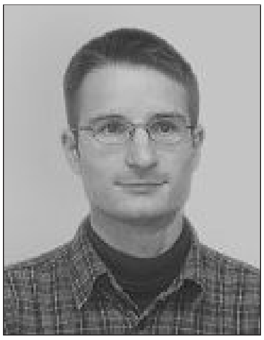
[23] D. Wiegand, "Mobiles facility management – Forschungsprojekt MUMMY", *Immob. Busin.*, vol. 4, p. 53, 2004.

[24] T. Hoffmann and T. Schwarz, "Digitale Datenerfassung im Feld", *ARCADIS Forum*, no. 17, 2004.

[25] H. Jacob and T. Hoffmann, "Mobiles Wissensmanagement: das Projekt MUMMY", *ARCADIS Forum*, no. 17, 2004.

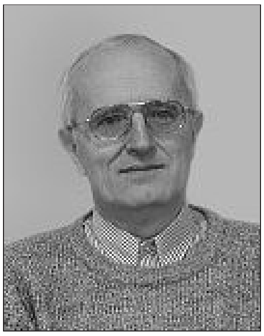
Martin Klima received the Dipl.-Ing. degree from Czech Technical University in Prague, Czech Republic, and in 2000 he has began his Ph.D. studies in the same university. In 2001–2002, he was working on the adaptation of multimedia documents formats to PDA at the ZGDV in Darmstadt, Germany. Since 2004 he is working in the area of usability testing especially with special regard to mobile devices. His fields of interest are: mobile user interaction, data adaptation for mobile environment, and collaborative work in mobile environment.

e-mail: xklima@fel.cvut.cz
Department of Computer Science and Engineering
Czech Technical University in Prague
Karlovo nam. 13
Praha 2, Czech Republic



Zdenek Mikovec received the Dipl.-Ing. degree from the Czech Technical University in Prague, Czech Republic, and in 2000 he has began his Ph.D. studies in the same university. In 2001–2002 he was working on the adaptation of multimedia documents formats to PDA at the ZGDV in Darmstadt, Germany (analysis, design, prototype implementation). In 2005, he was an Assistant in the Department of Computer Science and Engineering at the Czech Technical University in Prague. His fields of interest are: formal picture description, special user interfaces and XML/XSL.

xmikovec@fel.cvut.cz
Department of Computer Science and Engineering
Czech Technical University in Prague
Karlovo nam. 13
Praha 2, Czech Republic



Pavel Slavik received the Dipl.-Ing. and Ph.D. degrees from the Czech Technical University in Prague, Czech Republic, in 1970 and 1983, respectively. In 1994, he was an Associate Professor and in 2003 Full Professor, in the same university. His fields of interest are: computer graphics, scientific visualization, graphical user interfaces and interfaces for users with special needs.

slavik@fel.cvut.cz
Department of Computer Science and Engineering
Czech Technical University in Prague
Karlovo nam. 13
Praha 2, Czech Republic

Dirk Balfanz graduated with a degree in electrical engineering from the University of Dortmund, Germany, in 1992. From 1993–1996, he worked in the area of digital wireless communicationat BOSCH Telecom and specifically implemented one of the first DECT-based systems. In 1996, he joined the Fraunhofer Institute for Computer Graphics as a researcher in the field of metadata technology for geographic information systems. Since 2000, Doctor Balfanz has been Head of the Research Department of Mobile Information Visualization at the Computer Graphics Centre (ZGDV e.V.) in Darmstadt. As the leader of a team of computer scientists and engineers, he is responsible for research and development in the area of context-sensitive multimedia data processing and visualisation in mobile systems. Current research efforts focus on the development of situation- and context-aware mobile applications with multi-user collaboration and mobile video interaction.

e-mail: Dirk.Balfanz@zgdv.de
ZGDV – Computer Graphics Center
Rundeturmstrasse 6
64283 Darmstadt, Germany

Paper

Empowering the mobile worker by wearable computing – wearIT@work

Michael Boronowsky, Otthein Herzog, Peter Knackfuß, and Michael Lawo

Abstract— Currently wearable computing is still a technology of niches and in a laboratory stage. With wearIT@work a project dedicated to applications was launched by the European Commission (EC IP 004216). The first year of the project is nearly over and first results were achieved. In this paper the concept of the project is introduced and first results are presented. As the project strongly follows a user centred design approach much effort was put on first investigations with users in the four application domains of maintenance, production, hospital and fire fighting. Beside this results concerning a wearable computing hardware platform and software framework were achieved.

Keywords— wearable computing, applications, user centred design.

1. Introduction

The European Commission set up an *Integrated Project* wearIT@work [1] to investigate wearable computing as a technology dealing with computer systems integrated in clothing. The project has 36 partners, among them EADS, HP, Microsoft, SAP, Sony, Siemens, and Zeiss. With a project volume of 23.7 million € and a funding of 14.6 million € under contract no. 004216, wearIT@work is the largest project world-wide in wearable computing. The TZI – Mobile Research Center of the University of Bremen coordinates the project.

The project wearIT@work contributes to the shaping of today's most challenging computer applications. The intention of wearIT@work is to prove the applicability of computer systems integrated to clothes, the so-called wearables, in various industrial environments. These novel computer systems support their users or groups of users in an unobtrusive way wearing them as a computer-belt. This allows them to perform their primary task without distracting their attention enabling computer applications in novel fields. Interaction with wearables by the user is minimal to realize optimal system behaviour. For this reason a wearable computer has to recognize by integrated sensors the current work progress of a user. Based on the work context detected the system has to push useful information to its user, e.g., how to proceed with the work. Apart from speech output, media could be optical systems presenting the information, e.g., via semi-transparent glasses within the worker's visual field. Output devices for tactile feedback are also applicable.

One of the major goals of the project is to investigate the user acceptance of wearables. Suitable methods for user interaction and processes suited to wearables in industry are identified. Investigations show that methods to detect the work context and a general architecture of wearables as well as a hardware and software platform for the implementation of wearables are urgently needed. Four industrial pilot applications, namely emergency, variant production, maintenance, and the clinical pathway drive the project.

The focus of the emergency activity field is the collaborative planning and interaction using wearable devices. In variant production the challenge is the information integration and the intelligent information presentation. The maintenance scenario has its focal point on context detection and intelligent manuals. For the clinical pathway the focus is on intelligent information logistics and context aware collaboration.

The *Integrated Project* is organized in activity lines (AL) and activity fields to manage its complex structure and follows a user centred design (UCD) approach (Fig. 1). Applications are developed in cycles of 18 months duration each.

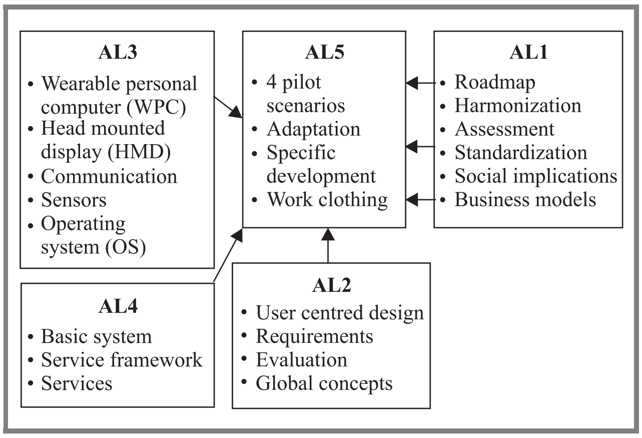


Fig. 1. Organization of wearIT@work.

The partners of the project go for the following advantages. First of all, an improved productivity and flexibility of workers shall be reached. Second, an increased safety at work and a decreased pressure towards automation is aimed at. Third, a simplified access to enterprise information, and fourth, faster group decisions are intended. Last but not least, new information technology products will be introduced into the market based on the pilot applications developed within the project.

"*The worldwide market for wearable computers generated over $70 million in supplier revenues in 2001. The market will increase at a compound annual growth rate (CAGR) of over 51% through 2006, and grow to over $563 million*" [1].

Despite its massive growth, the market for wearables is still a niche market compared to the industrial use of desktop computers. Drivers of a stronger growth will be more standardized hardware and software platforms enabling the new work paradigms.

2. Relation to existing theories and work

There are different approaches to defining wearable computing depending on the research direction and the application domain. In the wearIT@work project the interaction between the user, the system and the environment is focused on. In conventional mobile systems the interaction is based on a modified version of a desktop human computer interface (HCI) and follows the pattern shown in Fig. 2a.

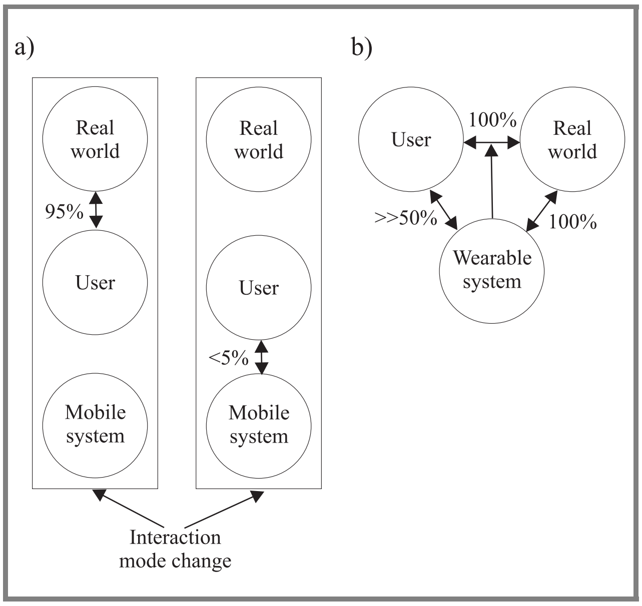


Fig. 2. Interaction between the user, the system and the environment in a conventional mobile system (a) and a wearable system (b).

To operate the system, the user needs to focus on the interface. This catches his attention as well as his physical activity, in particular the use of his hands. As a consequence he/she can either interact with the system or with the environment, however not with both of them at the same time.

Thus to access data on a personal digital assistant (PDA) the user must interrupt whatever he is doing, take the PDA out of his pocket and focus on the device. This mode of operation implies that the range of applications and situations in which the system is useful is severely restricted.

In general, over a course of a day, mobile devices are actually useful at most between 1% and 5% of the time. In addition, many of today's devices are bulky and obtrusive which means that for many activities the user does not even take them along.

In contrast, wearable systems are designed to be permanently useful and usable in a wide range of mobile settings. The corresponding interaction concept is illustrated in Fig. 2b. It allows the user to simultaneously interact with the system and the environment. In addition, there is direct interaction between the system and the environment as well as the possibility of the system mediating the interaction between the user and the environment.

The implementation of the wearable interaction concepts involves four main issues:

1. The system must be able to interact with the environment through an array of different sensors distributed in different parts of the outfit. In particular it must be able to develop a certain degree of awareness of the user activity, his physiological and emotional state, and the situation around him. This is often referred to as **context awareness**.

2. The user interface needs to be operated with minimal cognitive effort and with no or little involvement of the hands. In general, the low cognitive load is achieved through appropriate use of the context information. Thus for example instead of having the user select a function from a complex hierarchy of menus, the system should derive the two most likely options from the context information and present the user with a simple binary choice. In terms of the actual input modality, simple, natural methods such as nod of the head, a simple gesture, or spoken commands are preferred.

3. Using context information the system should be able to perform a wide range of tasks without any user interaction at all. This includes system self-configuration tasks as well as automatic retrieval, delivery, and recording of information that might be relevant to the user in a specific situation. A trivial example of a context-dependent reconfiguration could be a mobile phone that automatically switches off the ringer during a meeting.

4. The system must be seamlessly integrated in the outfit so that it neither interferes with the user's physical activity nor affects his appearance in any unpleasant way. This means that unlike many conventional mobile devices, it can be taken along nearly anywhere.

A trivial example of an existing device that adheres to the above requirements is a modern hearing aid computer. It is unobtrusive, useful during most of the day, requires hardly any cognitive effort to operate, and by definition of its function, mediates the user's perception of the real world. In addition advanced devices are able to automatically adjust the volume between noisy and quiet settings and even optimize the amplification mode to suit the situation such as conversation or a concert.

In wearIT@work we address more advanced wearable systems. They can detect complex activities such as social interaction, or certain specific work related actions (e.g., in maintenance) and use this information to deliver a variety of services exactly tailored to the user needs in a given situation.

3. Research approach

The research is based on the user centred design approach as defined in ISO 13407 [5] (see Fig. 3). Based on scenario definitions and discussions between the stakeholders of the project and the application partners workplace studies were performed on the site of the users to validate the scenarios. In the next steps mock-up prototypes are developed and evaluated again with all stakeholders at the user sites.

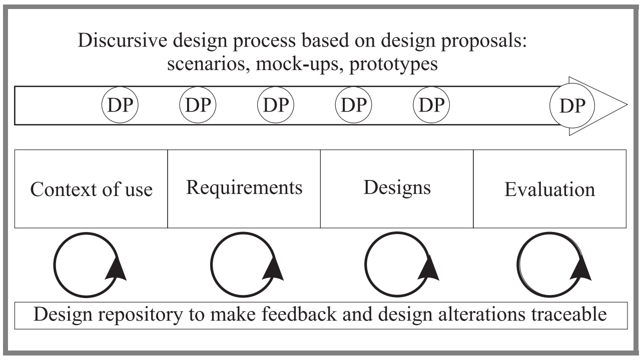


Fig. 3. Generic ISO 13407 compliant UCD process model.

The social computing factors are with wearable computers similar to those of laptops and mobile phones. Assumed wearable computers are provided by the organization to the employees, they might be used beyond normal working hours, beyond the physical borders of the organization, like at home or beyond the organizational work context like for private or public purposes not related to the employee's job. The results of these circumstances are many partially interacting factors as those given in Fig. 4.

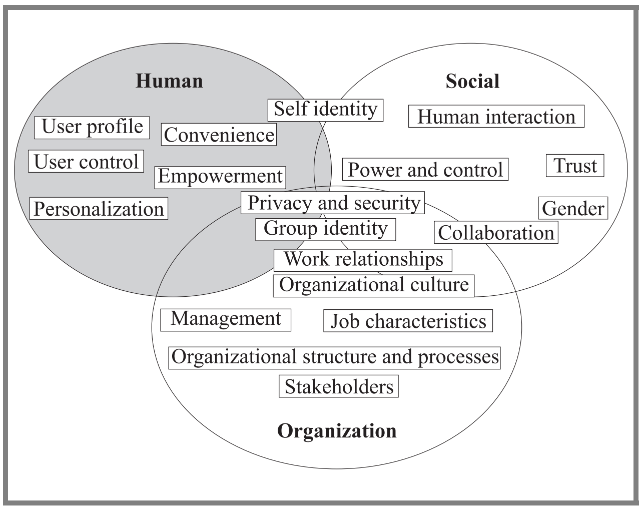


Fig. 4. Social computing factors map.

Aspects like self identity, privacy and security or group identity need a clear understanding and negotiations by the stakeholders for achieving a common understanding of the implications and consequences of the use of a wearable computing technology in a professional environment as addressed in the project.

4. Findings

Beside the application oriented interaction with the end users of the project a lot of effort was put into a common understanding of wearable computing (see above) and first steps towards a common wearable computing platform and framework.

A hardware platform consisting of a core wearable computing unit, input and output devices, general peripherals, and

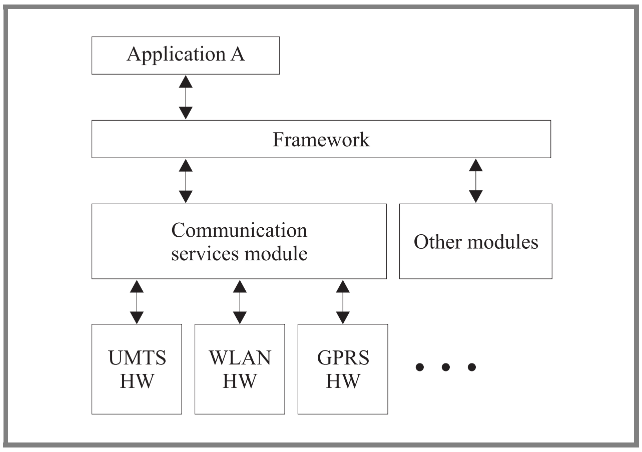


Fig. 5. General wearIT@work architecture.

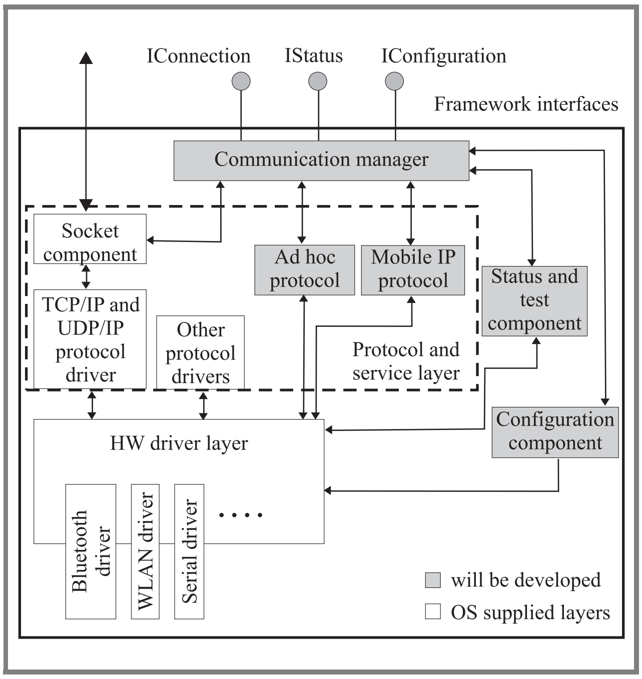


Fig. 6. Communication service module architecture.

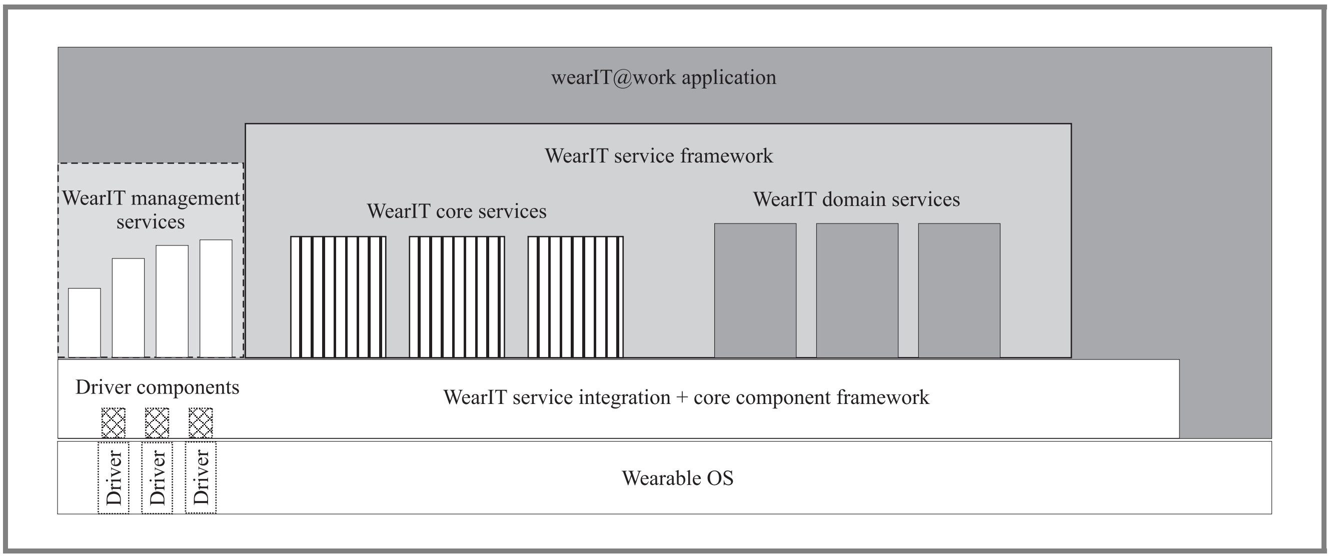


Fig. 7. General structure of the software framework.

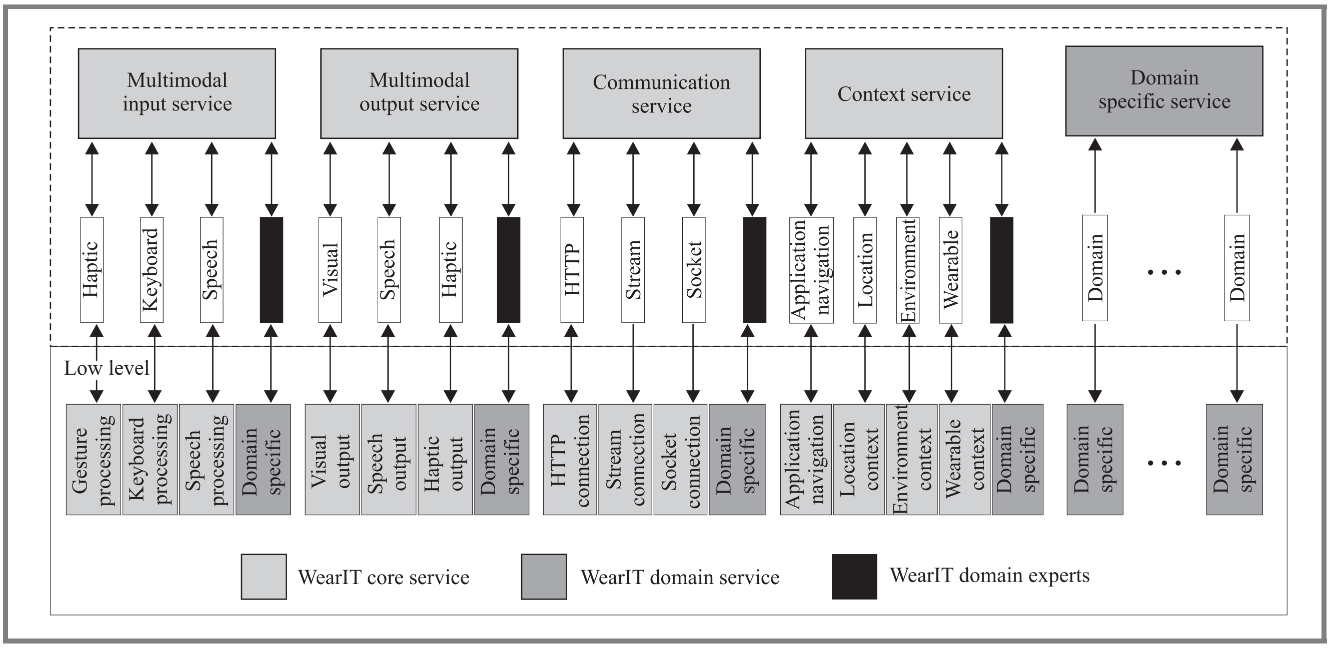


Fig. 8. Service framework.

sensor and communication subsystems was defined. One of the challenges is the necessity to provide the end-user with a seamless access to heterogeneous networks. This reflects the general wearIT@work architecture as shown in Fig. 5 as well as the communication service module architecture as given in Fig. 6.

The idea of creating a common software framework based on a common hardware platform is of great importance from the perspective of the project as well as beyond. Only in the case that wearIT@work will be successful with this process a remarkable impact is achieved for the exploitation of wearable computing solutions. In this case not only the four domains addressed within the project application but also other related application domains will benefit from the result and in the case of the addressed standardization push also the developers of devices, components and solutions will be covered.

The general structure of the software framework covers beside a service registry and high level services also core services like context awareness, communication, I/O, and security (see Fig. 7). Beside these core services domain specific services of the same structure exist. The idea is to integrate services of common use within the application domains of the project into the core services. The advantage of this approach is that with increasing and/or changing requirements the general structure remains valid. Figure 8 illustrates this fact; core and domain services can be made available on different levels.

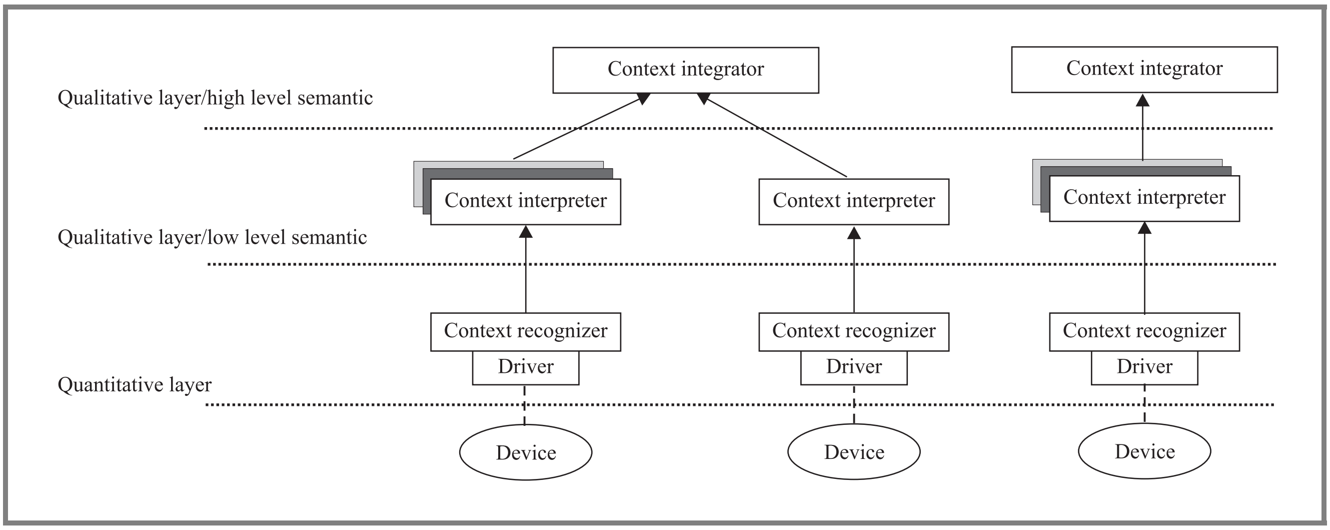


Fig. 9. Context detection using sensors.

Context detection is one of the essential success factors of wearable computing as mentioned above. Only in the case of using sensors a context can be detected with high reliability, and the cognitive load of the end-user can be managed in a successful way. This is considered to be one of the most scientifically challenging topics of the project. Other challenges like the robustness required always outside the lab in the real world are more relevant for producers and developers of devices.

To master the context detection problem, a general approach was discussed which is suitable for extension and adaptation (see Fig. 9). It is foreseen to perform series of tests with the end-users using existing sensor subsystems that are still basically in a prototype stage. These end-user tests are necessary to decide which further research and development work on sensor subsystems is necessary to achieve the performance accepted by the end-users.

5. Conclusion

In this paper the results of the first nine months of wearIT@work were presented. There are still nearly four years of research to be done and there are still many results to be achieved, but the fundamental steps towards a user centred design approach, a hardware framework and software platform are done. With the creation of the Open Wearable Computing Group [3] and the International Forum on Applied Wearable Computing [4] organized annually, a community building process in industry and science has been initiated. It is the intention of the project and the accompanying activities to understand the project as the leverage for wearable computing, even if miniaturization and low power computing devices as well as ubiquitous wireless communication are still in an emerging stage to provide the wide spectrum of innovative solutions which is necessary to achieve wearable computing anytime, at any place, and in any situation.

Acknowledgement

This work has been partly funded by the European Commission through IST Project *wearIT@work: Empowering the Mobile Worker by Wearable Computing* (no. IP 004216-2004). The authors wish to acknowledge the Commission for their support. We acknowledge our gratitude and appreciation to all the 36 wearIT@work project partners for their fruitful work and contribution in the research work done so far and forthcoming.

References

[1] " wearIT@work", http://www.wearitatwork.com/

[2] T. Shea, "The global market for wearable computers: the quest for killer applications", Venture Development Corporation, Aug. 2002.

[3] M. Lawo *et al.*, "The Open Wearable Computing Group (OWCG) – a concept for standardization", in *Proc. 2nd Int. Forum Appl. Wearab. Comp.*, Zurich, Switzerland, 2005, pp. 181–182.

[4] "International Forum on Applied Wearable Computing", www.ifawc.org

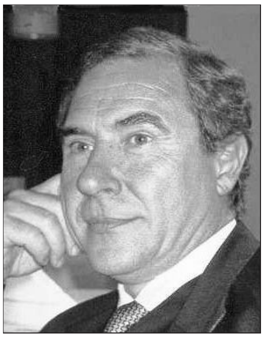
[5] "Human-centred design processes for interactive systems", ISO 13407, 1999.

Michael Boronowsky is Managing Director of the Center for Computing Technologies (TZI), University of Bremen, Germany. In 1992 he had finished his studies in electrical engineering at University of Applied Science Aachen. After which he worked in a measurement device company as leading development engineer. He went back to university and studied computer science in Nijmegen, Netherlands, and received a Master in computer science in 1995. Since this time he is working at

the University of Bremen. In the beginning he was working in the area of artificial intelligence research. He developed machine learning algorithms, to learn qualitative physical models from the observation of measurement data. In his Ph.D. he focussed on methods for the automatic interpretation of time series data based on decision tree induction. In 1999 he became responsible in co-ordinating the wearable computing research at the TZI. He finished his Ph.D. in 2001 and became Managing Director of the Intelligent Systems Department. Since 2002 he is Managing Director of the TZI. He coordinated the application of the IP-proposal "wearIT@work – empowering the mobile worker with wearable computing" featuring a groundbreaking vision of the computer empowered knowledge worker of the future. Doctor Boronowsky is Vice Chair of the European Community "mobility@work" and also Vice Chair of the overall AMI@work community focusing on the economical breakthrough of the ambient intelligence vision for the mobile worker. Beside his management activities he is still working as senior scientist in research projects related to mobile computing and is supervisor of several master thesis in this area.
e-mail: mb@tzi.de
Center for Computing Technologies (TZI)
University of Bremen
P.O. Box: 33 04 40
D-28334 Bremen, Germany

Peter Knackfuß was born in 1955 in Munich. After school he went to Fachhochschule Munich and to University of Munich, where he left 1984 with German Diplom degrees in mechanical engineering and in electrical engineering. Then he joined Gesellschaft für Reaktorsicherheit in Munich working as a software engineer. In 1987 he changed to the Bremen Institute for Industrial Engineering and Work Science – BIBA starting as a research engineer and working as a department leader later on. In parallel he wrote Ph.D. thesis at University in Bremen and was graduated as a doctor engineer in 1992. In 1997 he left BIBA and founded his own company dealing with network technology and with support of industry and governments to enter European funding schemes. Doctor Knackfuß has continuously performed work in more than 15 European projects as a researcher since 1984 and has till now managed about 10 European research projects in different European funding programmes.
e-mail: Peter.Knackfuss@infoconsult.nu
Info Consult GmbH
University Avenue 17
28359 Bremen, Germany

Otthein Herzog holds a chair in artificial intelligence in the Department of Mathematics and Computer Science at the University of Bremen, Germany. He directs the TZI – Center for Computing Technologies, the Collaborative Research Center 637 – Cooperating Autonomous Logistics Processes sponsored by the Deutsche Forschungsgemeinschaft, and the Mobile Research Center in the Mobile Solution Center Bremen. His current research interests include wearable and mobile computing, multi-agent systems, semantic data transformation and mediation, content-based analysis and retrieval methods for digital libraries, and digital image processing. Before Doctor Herzog joined the University of Bremen, he worked for IBM in various R&D and managerial positions for 16 years. Professor Herzog has co-authored and co-edited seventeen books and more than 80 scientific papers. He is a member of Gesellschaft für Informatik, DAGM, ACM, and the IEEE Computer Society.
e-mail: herzog@tzi.de
Center for Computing Technologies (TZI)
University of Bremen
P.O. Box: 33 04 40
D-28334 Bremen, Germany



Michael Lawo is with Center for Computing Technologies (TZI) at the University of Bremen since 2004 as the Technical Manager of the IP 2004-004216 wearIT@work after being the CEO of a group of SME in the IT domain since 1999; from 1996 to 2000 he was CEO of an IT consulting firm which now belongs to UNILOG S.A./Paris and from 1991 to 1995 top manager information systems with the Bremer Vulkan group. He was consultant before joining the nuclear research centre in Karlsruhe from 1987 to 1991 as head of the industrial robotics department. He is a 1975 graduate of structural engineering of Ruhr Universität Bochum, received his Ph.D. from Universität Essen in 1981 and became Professor in structural optimization there in 1992. In 2000 he was appointed as Professor of honour of the Harbin/China College of Administration & Management. He is author, co-author and co-publisher of five books and more than 110 papers on numerical methods and computer applications, optimization, IT-security and wearable computing.
e-mail: mlawo@tzi.de
Center for Computing Technologies (TZI)
University of Bremen
P.O. Box: 33 04 40
D-28334 Bremen, Germany

Paper

EWD-P as an example of a "The Best of Good Practice" project

Grzegorz Bliźniuk, Dorota Biernacka, Grażyna Kowalska, and Mariusz Momotko

Abstract— The European Document Exchange System – Poland (EWD-P), developed by Rodan Systems SA is one of the most modern electronic document exchange systems in European Union. EWD-P is a workflow system that facilitates decision-making within the state bureaucratic system, with a special brief to work out official Polish government standpoints on numerous legislative issues constantly arising within the EU. The EWD-P project has created an effective platform for electronic exchange of documents related to the EU legislative process. The intelligent workflow management functionality aims to support a complex flow of documents through the meanders of the central government administration. The EWD-P system includes a high-level classification of documents using an artificial intelligence technique for document categorization. The EWD-P system simplifies interaction between ministerial departments involved in elaboration of a final common position and facilitates a more efficient organization of work within government and other public administration institutions involved in the EU legislative process.

Keywords— The European Document Exchange System – Poland, EWD-P, workflow engine, classification and categorization of documents, central repository, advanced search engine, Rodan Systems, best practice project.

1. Introduction

1.1. Specific problem

An important ingredient of the European Union enlargement process was to ensure conformity of a candidate state's existing legislation to the EU procedures and regulations. Such conformity, but with reference to proposed new legislation continues to be required after the accession.

Elaboration of an official Polish standpoint is a complex and recurring process (more than 2000 process instances a month). It requires involvement of relevant units of central government (all ministerial departments) and collaborative efforts of selected advisory groups from various bureaucratic domains (the state employs more that 12 000 civil servants). The results' delivery within often tight time limits is critical because failure to deliver on time means unconditional acceptance of the proposed EU directive. The official Polish government reply to a EU proposal has at times to be passed on to the Polish Representation Offices at Brussels in a very short time – sometimes counted in hours rather than days.

Such blitz-like decision-making requires high-level techniques of classification and document distribution as well efficient cooperation between various teams of advisors. Before the introduction of a dedicated IT system a significant percent of standpoints could not be properly elaborated, hampering the overall integration process [1].

In pre-EWD-P days, the main problem concerned expensive and time consuming procedures in distribution of EU documents to Polish institutions and experts. There were also considerable difficulties in accessing the central repository and thus the history of the official standpoint elaboration process. In short: there were problems in the process of identification and procurement of information as well as of related documents (each institution had different databases that were not synchronized). Together these difficulties prevented the government from applying an optimal and proper standpoint elaboration process and underlined the noticeable gap in the decision-making processes between the internal (interdepartmental) and external (Poland–UE) standards.

1.2. General background

The European legislative procedures are based on a complex infrastructure of working groups and committees focusing on assorted domains governed by the European Parliament regulations and the European Commission directives. Representatives of all member states participate in the legislative process preparing appropriate position documents and participating in respective working group and committee meetings. In order to facilitate the document exchange between the General Secretariat and member state governments an information system called the U32 mail was implemented in 2002. Each member state (receiving about 100–300 documents daily) has been provided with two access points to the U32 mail system, one in a selected location in the member state central government and one in the Brussels representation office. Most of the documents pertain to new issues requiring a position to be presented by the member state government, the rest pertain to the already existing, i.e., "current", cases. The U32 mail system initially classifies the documents and enables defining simple distribution rules.

For the majority of the EU documents a Polish response must be prepared and delivered by a certain date. If the deadline for preparation is not met, it is assumed that there is no objection to the proposals contained in the original EU document, which is deemed fully accepted. When a Polish response is attained, it is sent to the Representation Office at Brussels which in turn presents this reply to the Council's relevant office.

Stringent discipline imposed by the European legislative procedures, usually involving several central government agencies and authorized individuals participating in various auxiliary roles, calls for advanced IT functions to provide the required level of technological support.

The EWD-P system enlarges the powers of the U32 mail system in a significant way. It supports the process of elaboration of Poland's official government standpoint and stores all documents send via U32 mail from General Secretary of the Council of the European Union and form the country that at the time presides over the EU Council.

1.3. Policy context and strategy

In 2004 Poland became a member of the European Union. From that date on, the process of elaboration of the official Polish standpoint is governed by a directive requiring obligatory consultations of Poland's government standpoint between the ministerial departments.

Poland's IT strategy – ePoland 2004–2006, defining the direction of IT development for the public administration has been prepared by the Polish government.

The EWD-P system has been developed and implemented to provide an effective platform for electronic exchange of documents involving the UE legislative process between all Polish public institutions.

2. Solution

2.1. Objectives

The prime objective was to create a system, which would assure that the complex process of elaboration of a Polish government standpoint is executed efficiently, accurately and, most importantly, on time.

More specific objectives:

- Storing in one place a full knowledge related to subjects concerning UE legislative process and Polish public institution.
- Monitoring a progress of tasks within a process of working out a standpoint for a specific case.
- A selection of proper experts for specific cases and possibility of consultation among them.
- Accessing a full history of the elaboration process for a specific case.

2.2. The project

In order to achieve these objectives a project called EWD-P was launched with responsibility for creation of a system that enhances the individual advisors' and experts' input in the process of elaboration of the standpoints and improves collaboration during the meetings of various working groups.

The range of tasks within a scope of the EWD-P project included:

– system analysis;

– conceptual design of the system;

– technical design of the system's components;

– development of the application;

– testing and verification;

– procurement and ddelivery of equipment (hardware, including servers and data communication devices);

– delivery of software licences required for system development and operations;

– setting up a production and a training environment (installation of all software and the application);

– import of historical data;

– preparation of system's documentation;

– organization and execution of training for ca. 1000 users divided into three groups: coordinators, departmental coordinators and experts;

– technical support for the system users; the EWD-P system has to meet the needs of users from 19 government ministries and institutions.

The duration of the project was very short for such a complex solution. The EWD-P system implementation and deployment was possible only due to the fact that the project used results constructed within a R&D project – ICONS (IST-2001-32429) [3]. The development was based on generic and powerful services of the ICONS platform for knowledge management intensive portals development. The EWD-P system is the first commercial exploitation of the ICONS platform.

2.3. The EWD-P system

The system has been designed to apply technology directly in support of business processes. The EWD-P application has been developed in OfficeObjects®Workflow environment and J2EE technology. The architecture of the system is illustrated in Fig. 1.

The main features of the system:

- A central repository armed with a document management mechanism:
  - automatic reading of information from U32 mail system;
  - classification/indexation of documents;
  - cross-reference between documents;

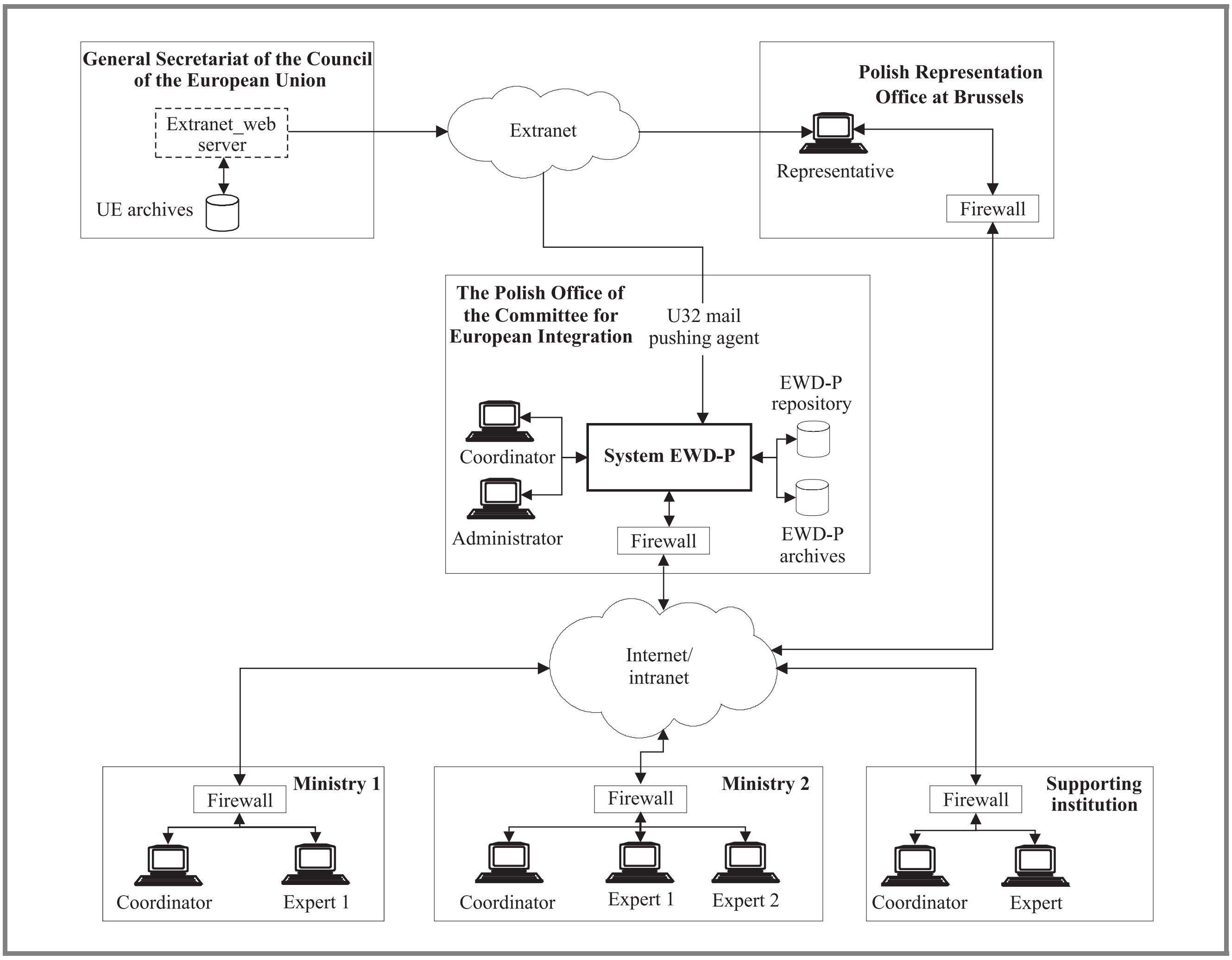


Fig. 1. The EWD-P system architecture.

  – pre-defined categories of documents and cases (according to acronyms, authors/persons engaged, working groups);
  – storing versions of documents;
  – structural retrieval (using attributes) and full text retrieval.

- Workflow coordination:
  – definition of workflow processes of elaboration of standpoints and instructions;
  – providing a list of users’ tasks.
- Diary of meetings of working groups, COREPER and the Council of EU:
  – automatic registration of meetings triggered by a note from the U32 mail system;
  – integration of the diary with cases and documents and the process of elaboration.
- Configuration of the system – meta-information management:
  – management of Polish administration structure management;
  – providing the EU and Polish topical classification of cases;
  – registration of advisors’ certificates and domains of interest;
  – users’ authorizations management.
- Notification module:
  – generating automatic notification of delays and other information relevant for persons engaged in a particular case;
  – sending and receiving notifications to/from other system’s users.
- Connections module.
- Help module.
- Personalization:
  – adjustment of views to a user’s role;
  – setting-up a user’s preferences.
- Replacement services.
- Web access.

2.4. Implementation

The most exciting aspect of this implementation was the potential it opened up for the convergence of a number of new technologies. Achieving compatible operation with the European Commission General Secretariat required providing solutions to several advanced technological problems including the use of artificial intelligence techniques for document categorization as well as an intelligent workflow management functionality to support complex documents flow through the quagmire of the central government administration.
To satisfy the above goals required sophisticated skills in designing and programming area coupled with deep knowledge of the subject.
The EWD-P system has been split into a number of modules. A short description of the main modules is presented below.

2.4.1. Notification

This module provides for automatic notification of system users regarding delays in completion of a allocated task in the process of position elaboration. The module also notifies the user when an important situation arises in the elaboration process (e.g., appearance of a new version of a document). The user is notified of crucial incidents effecting decision-making process (change of dates of meetings, projected absence of crucial users, etc.). The module allows sending notifications to other system users. Notifications can own notes to documents and cases placed in system.

2.4.2. Tasks

This module is responsible for steering the process of working out a response to received EU documents. Based on metainformation stored in the system, a semi-automatic or automatic classification of these documents is performed. Documents are distributed to relevant departments and expert advisors. A group of experts prepares a document of position or instruction, which after acceptance, is passed to the Polish representative at the EU in Brussels. The representative then presents such worked-out position during a workgroup or COREPER committee meeting. The minutes indicating the most important issues discussed during these meetings are registered in the system. The report, containing feedback from cases discussed at the meeting, is made available to system users, including the advisory group, which elaborated the original position.

2.4.3. Supporting cooperation between advisors

This module is directly connected with the tasks module. It supports and records the flow of discussion between advisors taking part in the decision-making process. Experts can present point of view of departments, which they represent, in the form of documents registered on a given forum. This information is confidential and is made selectively available to a working group directly involved in elaborating the position for a given case.

2.4.4. Documents

This module is responsible for a management of documents received via U32 mail or registered manually and also documents created during the process of elaboration (standpoints, instructions, other documents).
The module allows:

– registration of documents send by U32 nail system, which are coming to a dedicated e-mail address;

– addition and editing of new documents (description attributes and files-attachments);

– publication of documents (making them available to other users);

– creating and recording new version of an existing document;

– documents searching using attributes (e.g., type, language) and full text searching.

2.4.5. Cases

A case is created when a new document is registered, which is not related to an existing case. Otherwise the document is added to the existing case. The case is closed automatically once a position has been worked-out.
The module allows browsing of all cases registered in system. Cases can be searched by a number, description, subject or by a name of a person connected with the case. Any discussion related to this case might be looked at as well. Also a history of process realization can be examined (process chart and activity list in process).

2.4.6. Diary

This module allows browsing of information gathered in the system about working group meetings on which the elaborated standpoints are presented. The information on these meetings are taken from messages coming via U32 mail system and from coordinators, experts or the Polish Representation Office at Brussels. The module allows also registration of an agenda and linking it with an appropriate meeting. This is done automatically by the system or manually if the source of the document containing an agenda is outside the U32 mail system. Users are able to register other documents related to the particular meeting (e.g., an instruction for Polish representatives, reports). Meetings can be searched by attributes (e.g., meeting's code, EU working group's code, location, date, participators, discussed cases, etc.).

2.4.7. Configuration

This module coordinates a task of configuration management and enables definition of the metadata which is necessary for a proper working of the system such as: users, roles, organization units, working groups, UE subjects, Polish subjects and other dictionaries. It also defines various cross references between the metadata and sets other parameters needed by the system. The module allows browsing, editing, activation and deactivation of these elements and facilitates setting up dependencies between them. The system creates a full history of any change to the configuration.

2.5. Workflow engine

The EWD-P system dynamics is imposed by processes executed by the workflow engine – one process instance for each group of EU documents on one meeting. The workflow starts from publication of an original document on the U32 mail gateway, continues through initial dispatching in the Committee for European Integration, precise dispatching on the level of particular ministries' units (government's departments), establishing collaboration among the selected experts, elaborating of a final standpoint and ends up with the final delivery to the Polish Representation Office at Brussels. The Polish response preparation activities are supervised by a leading expert. At the end of this stage, all involved experts vote on the approval of the final Polish standpoint in the given case.

A single generalized workflow process type has been sufficiently powerful to support flows pertaining to any of the individual case categories involving a host of distinct agencies and individuals invoked in the appropriate roles to provide indispensable expertise. Advanced time modelling techniques have been employed to avoid delays in providing government position papers that might result in default official consent with respect to a particular issue.

Electronic records of government proceedings including all official and intermediate case documents as well as the workflow process information are maintained in the one central repository based on relative database, to provide invaluable reference material for government analysts.

Controlled access to the EWD-P system is provided for all participants via the open Internet with security supported by an elaborated role-based user access right model and the SSL data encryption. Electronic signature is available to support document authorization whenever required.

Each incoming EU **document is categorized along the predefined EU classification system**. Categories are just multiple values in a given attribute of the document's "envelope" – characteristics of the document as nature, source and destination. This set of categories determines receivers of the document in the Polish administration. The problem is that the Polish and EU classification systems are not unified yet and vary both in the sense of languages and in the sense of the categorization power. One EU category can be mapped to several Polish categories. Similarly, one Polish category can be mapped to several EU categories. The approach applied in the EWD-P system is based on ontology implemented as the topic map. The EWD-P ontology, constituting the logical heart of the system, provides mapping between inconsistent set of concepts [1].

The EWD-P system ontology **manages information on each civil servant or government organizational unit** (represented as topics) that are involved in the standpoint elaboration process. There are complex relationships among experts, units and categories cover, e.g., person and organizational unit (is employed in), person and category (is an expert in), organizational unit and category (is responsible for). The relationships, modelled as associations and exploited by the workflow engine, allow for precise selection of the most suitable (knowledgeable) experts and the responsible unit.

Finally, the EWD-P ontology, although serving various purposes (out of which interoperability is the most important), is internally consistent and play the vital role in the end users' education and communication.

The implementation of the EWD-P dynamics is fully based on the workflow management technology. The workflow process explicitly defined in the form of conceptual business process model and XML process definition language (XPDL) is **responsible for classification and distribution of the EU official documents as well as for preparation of the Polish response**. The process of elaboration of the Polish standpoint is instantiated for every incoming EU document and then executed by the workflow management system.

The major challenge addressed by the workflow engine in the EWD P system was to optimize workflow processes executions in order to "**to assure that appropriate activities will be performed by the right (knowledgeable) participants based on pertinent information in due time**".

A process instance is perceived by the workflow participant through a task list in which all commitments of a given user (possibly involved in a numerous cases) are sorted along with some priorities.

The **one central repository** based on relative database includes full collection of information concerning all the standpoints' elaboration process.

3. The good practice

3.1. Impact

The obvious benefits of the implementation of EWD-P system are as follow:

- Reduction of time of the documents distribution to coordinators and experts, in turn experts have more time for elaboration of the Polish position.
- Expansion of U32 mail system capabilities in classification and distribution of EU documents.
- Improvement and automation of classification and distribution of EU and Polish documents.

- Provision of proper instruments supporting the process of creation of Polish topics' classification and design of a method of establishing a relation between Polish and EU topics.
- Provision of mechanisms supporting communication between departments and various experts in the process of Polish standpoints elaboration.
- Possibility of monitoring progress of elaboration process for each case.
- Automatic warnings by the system of unusual situations such as delays, possibility of missed deadlines, etc.
- Opportunity for on-line collaboration between experts and swift resolution of arguable problems.
- Availability of refined search facility (cases and documents by their attributes, content, and history of execution).
- Enabling access to historical information.
- Management of proficient document exchange among Polish institutions.

The EWD-P system simplifies interactions between Polish government departments (ministries) working on elaboration of the common state. Owing to implementation of the EWD-P system, the several important changes have taken place in a approach of government's departments to UE legislative process. The first homogenous, over-departmental topics' classification was introduced. It facilitates a better organization of work in ministries and other public administrations institutions that are involved in UE legislative process. The one mutual process of elaboration of the common standpoint has been defined and implemented. That is uniform process for all users across the government's departments. Due to implementation of the EWD-P system the Polish experts have instant access to key information concerning Polish standpoint's elaboration process. Moreover, owing to workgroup platform and collaboration facility they can communicate immediately among themselves to resolve any arising issues instantly, even if the experts are miles apart.

The EWD-P system has been successfully deployed in the Polish Office of the Committee for European Integration (UKIE) and all of Polish ministries. End users represent experts and coordinators from over 20 government's departments, central government agencies and other public institutions. At present the system has over 1000 of registered users. Daily documents' flow rate is 150 documents on average. The highly productive workflow-based environment ensures that the processes execution times for recorded cases vary from 2 to 14 days. Over 8 million pages of electronic documents are already stored in system repository.

Full deployment of EWD-P is planned to end in 2006. Over 10 000 end users cooperating with over 2000 working groups and numerous committees will take an advantage of the powerful functionality of the system.

3.2. Relevance of the case for other administrations that could learn from the experience

So far, the European Union has not worked out a common standard for collaboration work on EU procedures and regulations with newly associated states. Poland's example is one of the "case study" for our EU partners. The logic of the EWD-P system can be used as a template or a model for similar systems defining different processes, different topics and issues in inter-institutional and multi-organizational arena.

There are several features of the EWD-P system, that can be used in other administration's solutions:

- **The new way of process definition** based on a topics' classification.
- **Creation of the knowledge base of administration's processes** – the new procedures were defined in particular departments and institutions.
- **The EWD-P system could be a platform for new administration's services** – the possibility exists to create new processes, cases and documents.

The EWD-P system is great, implemented model of well planned complex administrative solution.

Demonstrated approach in this solution to document and process management has been acknowledged path to improve effectiveness and efficiency whilst saving valuable time.

3.3. Transferability

The EWD-P system can be easily transferred due to sophisticated metadata mechanism. There is a possibility of implementing EWD-P system in any country. To launch this system in other countries we have only to satisfy the following requirements:

- to define public administration's structure (ministries, departments, units, subsidiary institutions, etc.);
- to define a topics' structure (classification module);
- to define roles participating in the process;
- to register users and link them with a particular topic and previously defined role.

Experiences gained during realizations and practical usage of the EWD-P system, might be used and absorbed by any of a public administration institution, that demands execution of several related processes and presence of a central repository.

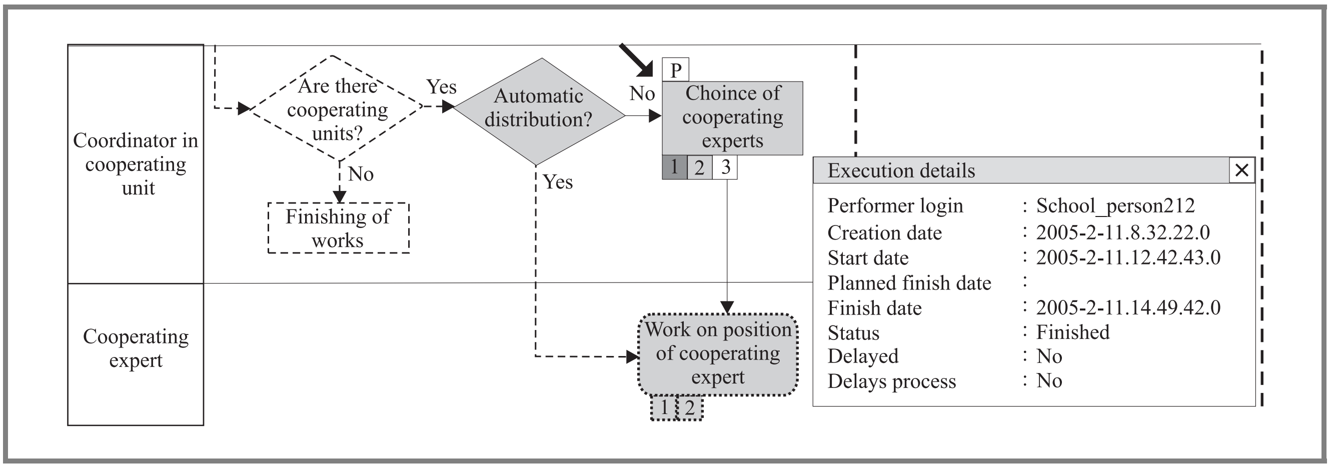


Fig. 2. An example of a process execution visualization.

4. Results

Benefits from the implementation of EWD-P system and work transformation of civil servants taking part in the Polish standpoint elaboration process is undisputable.
The EWD-P system is based on the workflow paradigm. The key results that made the workflow engine successful within the EWD-P system are presented below.

- **Assignment of workflow participants is dynamic**. The workflow engine uses the rules of assignment of experts and organizational units to documents’ categories according to the competence they have (carefully specified in the EWD-P ontology) or based on the role they already played in previous activities of a given process instance (e.g., let B will be carried out by the same person that carried out A). For instance, the Polish Ministry of Infrastructure is responsible for the “aviation” category. Let’s further assume that there is one coordinator (responsible on the ministry level) and two candidate leading experts in this ministry. The system suggests to the coordinator these two candidates and allows selection of the most suitable one (tacit knowledge application).

- **Communication among process participants is flexible**. Usually, the coordinator monitors progress of elaboration of the Polish standpoint and, if necessary, gives some hints and feedback to the experts involved to improve the quality. This communication pattern is difficult to express in the form of the traditional workflow process, since there is no algorithm defining when this communication occur and in which way it will be carried out. Thus, to complement strict rules of the traditional workflow processes, a mechanism of team collaboration management (TCM) has been introduced. This is important as especially teams creating new knowledge need more elastic forms of cooperation with more space left for innovation, creativity and spontaneity. Messages of various natures are grouped in thematic threads and moderated by the coordinator. Integration between workflow process and TCM assumes 2 dimensions: time and participants. The discussion forum is active from the elaboration process commencement up to the process termination (read only mode is available then). As participants are identified dynamically during the process enactment on the base of WPAL rules valuation, they are joined to the TCM as soon as they are selected by the workflow system (staring from the process owner during the process initialization). Since this moment he/she can read messages, start new threads and reply to messages of others. The discussion-forum-like facility fosters unconstrained collaboration of the process participants, what is really indispensable while complex, interdisciplinary and multidimensional problems are addressed.

- As Polish standpoints have to be prepared before some deadline, **time management** is of particular importance. If a standpoint is not prepared on time, the delivered EU document is accepted with no objection. Therefore, it is crucial to detect and signalize any delay in the process as soon as it occurs both on individual activity level as well as on the overall process level. To meet this requirement the ICONS workflow engine extends and implements time management algorithm proposed in [2]. This algorithm checks the deadline and duration constraints for the whole process and for its individual activities. It also determines the best and worst cases for these constraints taking into consideration different possible scenarios (paths) of the process execution. The predefined workflow participants (especially coordinators) are informed on delays. EWD-P system allows to select the optimal group of notified participants.

To mitigate the problems following from the standpoints elaboration process complexity (a lot of activities, a significant number of spread geographically organizational units and thousands of potential participants) a facility of **process execution visualization** in a human-readable form is used (Fig. 2). This graphical representation of the process instance (on the contrary to the visualization of the process definition) is an extension of the business process modelling notation. This visualization allows performers to better understand the process history (what was done before, by whom, what were the recommendations, what were the time constraints), presence (what its current state is, what are the requirements for the current activity) and future (who will continue the process, what are potential consequences of current decisions). As the work items (picked up from individuals' tasks list) are put into definitely broader context, it positively impacts participants' knowledge and consequently improves productivity and quality.

5. Conclusion

For the majority of the EU documents a Polish response must be prepared and delivered by a certain date. If the deadline for preparation is not met, it is assumed that there is no objection to the original EU document and that the document is fully accepted. When the Polish response is prepared, it is sent to the Polish Representation Office at Brussels which in turn presents the document to the council's pertinent teams.

From the information management point of view the "case" concept plays the crucial role as it secures all information related to a given response. This comprises the original EU document (triggering the corresponding case; perhaps in versions), response, related cases, external documents, etc. The cases are maintained by cooperating experts who attach the consecutive versions of the Polish response as well as other relevant information (e.g., legislative documents). The cases store minutes of meetings on which Polish representatives present the Polish response (this includes remarks and comments).

It is critical to ensure consistency with other cases and, to the same extent, preserve the information on previously existing cases that somehow influenced a given response. A net of relations among cases facilitates navigation among related cases and promotes the access to global view on some issues rather than to its individual aspects.

Besides its mission-critical functionalities the EWD P system serves a number of auxiliary services. First of all, the users can search over the repository of documents and cases using attribute-based search (supported by user friendly search criteria builder) and full text search. This covers also European documents repository, accessible via the U32 gateway. This fosters knowledge reuse based on previous experiences. Secondly, the system provides full accountability of experts' decisions by storing the information on contributions of individuals to the documents and cases as well as their behaviours in a concrete workflow process instance (e.g., missed deadlines). Thanks to this, in the case of any problems concerning the essentials of a given standpoint, the standpoint's development history can be examined in every detail. The complete information on cases (together with specific contributions provided by individuals) are archived according to the stringent central government practices and standards and accessible for search as well as retrieval but closed for any modifications.

The EWD-P solutions offer many immediate practical benefits, described in previous sections of this paper, as well as a platform for the future to help public institutions to keep pace with the remorseless rise in expectation.

The EWD-P system is at the forefront of developments in this area, which is still at the embryonic stage – but which will accelerate rapidly as the government's departments move towards a process based approach in their performance of daily tasks. Whatever the outcome of prepared Poland's IT strategy with reference to public administration is, today's EWD-P solutions have been designed and built to provide the systems foundations for tomorrow, delivering decision on vital issues when and where it is required.

References

[1] G. Blizniuk, M. Momotko, B. Nowicki, and J. Strychowski, "The EWD-P system Polish government – European Commission interoperability achieved", in *38th Hawaii Int. Conf. Comput. Sci. HICSS-38*, Honolulu, Hawaii, 2005.

[2] J. Eder and E. Paganos, "Managing time in workflow systems", in *Workflow Handbook 2001*, L. Fischer, Ed. Future Strategies, 2001.

[3] The IST-2001-32429 ICONS Consortium, "The ICONS software technical design manual", Aug. 2003, www.icons.rodan.pl

Grzegorz Bliźniuk was born on December 15th 1968. In 1992, graduated from the Cybernetics Faculty at the Military University of Technology (MUT) in Warsaw, Poland, with a Master's degree in cybernetics, specialized in information systems. In 2000, defended the doctoral dissertation on research methods for concurrent software quality before the Scientific Council of the Cybernetics Faculty at the Military University of Technology in Warsaw, receiving the Ph.D. title in technology, specialization: information systems. In June 2000, became manager of the Software Engineering Department at the Information Systems Institute of the Cybernetics Faculty at the Military University of Technology in Warsaw. On January 2nd 2003, appointed Director of the Department for Information Systems for the Public Administration in the Ministry of Science and Information Technology. He is an author

of the book “Badanie jakości programów współbieżnych” (in Polish).

e-mail: grzegorz.blizniuk@wat.edu.pl
Military University of Technology
Kaliskiego st 2
00-908 Warsaw, Poland

Dorota Biernacka received the M.Sc. degree in marketing and management from the Academy of Entrepreneurship and Management, Warsaw, Poland. Her professional life was connected with telecommunication and IT sectors, including mobile banking and merchandising. Since 2003 she has been worked on high technology marketing with a special focus on workflow and knowledge management topics. Experienced in dissemination of European Research Projects results. Currently she works as Marketing Director in Rodan Systems.

e-mail: dorota.biernacka@rodan.pl
Rodan Systems SA
Puławska st 465
02-844 Warsaw, Poland

Grażyna Kowalska graduated from the Department of Mathematics at University of Warsaw in 1976 with a M.Sc. in computer sciences. Her professional carrier took her through all aspects of IT from programming, systems analysis via project management to business management. Most of her professional life was spend working for a big multinational in the United Kingdom. In 1999 she became a Certified PRINCE2 Practitioner promoting this methodology in her work. Since she joined Rodan Systems as an Operational Director in 2004 she has been involved in various innovative projects culminating with her contribution to the article about the EWD-P system.

e-mail: grazyna.kowalska@rodan.pl
Rodan Systems SA
Puławska st 465
02-844 Warsaw, Poland

Mariusz Momotko is the Director of OfficeObjects® products department in Rodan Systems SA. He received his Ph.D. in computer science from the Institute of Computer Science of the Polish Academy of Sciences in 2005 (thesis about “Tools for monitoring workflow processes to support dynamic workflow changes”). He also received his M.Sc. in advanced computer science from University of Birmingham and M.Sc. in computer science from Technical University of Gdańsk. He has also been awarded Several TEMPUS sponsored scholarships. His interests are focused on business process management, workflow management and document management technologies. He is an author of several Polish as well as international publications on business process and workflow management systems. He has been participated in several EU founded research projects such as ICONS, COMPONENT+, ASG and OneStopGov. Doctor Momotko is the project manger and the chief architect of Polish part of the COMPONENT+, ASG, and OneStopGov projects.

e-mail: mariusz.momotko@rodan.pl
Rodan Systems SA
Puławska st 465
02-844 Warsaw, Poland

Paper

Innovative approach to identity management solution development for e-government at EU level

Kamelia Stefanova, Dorina Kabakchieva, and Lia Borthwick

Abstract— This paper presents the main aspects of research, analysis and design of the open identity management architecture for e-government development within GUIDE, a project financed by the 6FP of the EC. The most important identity management issues strongly influencing the European e-government development are briefly discussed. An emphasis is placed on the innovative interdisciplinary approach used in GUIDE, aimed at covering the whole range of technical, process, policy, legal and social identity management issues, and seeking to overcome the existing fragmentation of identity management initiatives. GUIDE brings together the European industrial, financial and technical market leaders in e-government solutions, as well as leading academic institutes of the relevant scientific disciplines. Through its scientific, technological and socio-economic goals GUIDE will contribute towards initiatives that will ultimately deliver multiple benefits to governments, citizens and businesses.

Keywords— identity management, open identity management architecture, e-government, interoperability.

1. Introduction

In March 2000, European heads of state and government meeting in Lisbon set the objective to make the European Union the most dynamic and competitive knowledge-based economy in the world by 2010, capable of sustainable economic growth with more and better jobs and greater social cohesion, and respect for the environment [1]. The e-government is one of the key areas of the EU's information society policy where further progress is required to reach the objectives of the Lisbon strategy. There is indeed a growing consensus that e-government is now becoming a key factor for increasing competitiveness. Better quality public services, more responsive and fit to their users' needs, provided electronically by more efficient public administrations, are perceived as essential to reap the benefits of the information society and reach the objectives of the Lisbon strategy.

For leaders in the public sector, the emerging debate over identity management and the selections of technology to authenticate citizens and business are among the most important of all matters to shape the information age advanced frames. The competing policy interests range from protecting citizen freedoms, privacy and other prerogatives on one end of the scale to ensuring law, order, national security and institutional efficiencies, on the other end.

Electronic identity management for e-government requires combination of technological, social, economic and application-oriented research, including security and privacy of the identity data; public trust and acceptability; technological, organizational and linguistic interoperability. The e-government at EU level needs a coherent approach for interrelations and compatible solutions.

This paper presents the main aspects of research, analysis and design of the open identity management architecture for e-government development within GUIDE[1]. The innovative approach towards the identity management issues within the e-government on a pan-European level, achieved as a result of the collaborative project partners' efforts, is briefly described.

2. Background

2.1. The identity management concept

Governments have always been concerned with identity and are now confronted by the unique challenge of provisioning identity in networked world. Managing identity is a fundamental piece of what a government does, and governments are vitally concerned with identity on a daily basis. Many of the lifecycle activities involved in creating, using, changing and ending an identity rest with governments. Electronic authentication and managing digital identities is certainly different in the government setting. Relationships between government, citizens and businesses are unique and may last a lifetime. Most importantly, individuals and organizations have higher expectations for government when it comes to protecting the privacy of information. There exists a tension between citizen and business demand for efficient and accountable e-government services and expectations for privacy protection.

Identity as a crucial security feature for e-government, can be considered as a uniquely defined and maintained set of data that refers to a person (natural or legal) and used for uniquely identifying the person for particular e-government processes. In other words, Identity depends on the range of government services in Europe. Identity management should support these processes in cross-application and cross border environments.

Going deeper in the concept of identity, we discover it is broad and complex. It is defined as the quality or condition of being the same, i.e., identity is what makes entities

[1] See, http://www.guide-project.org

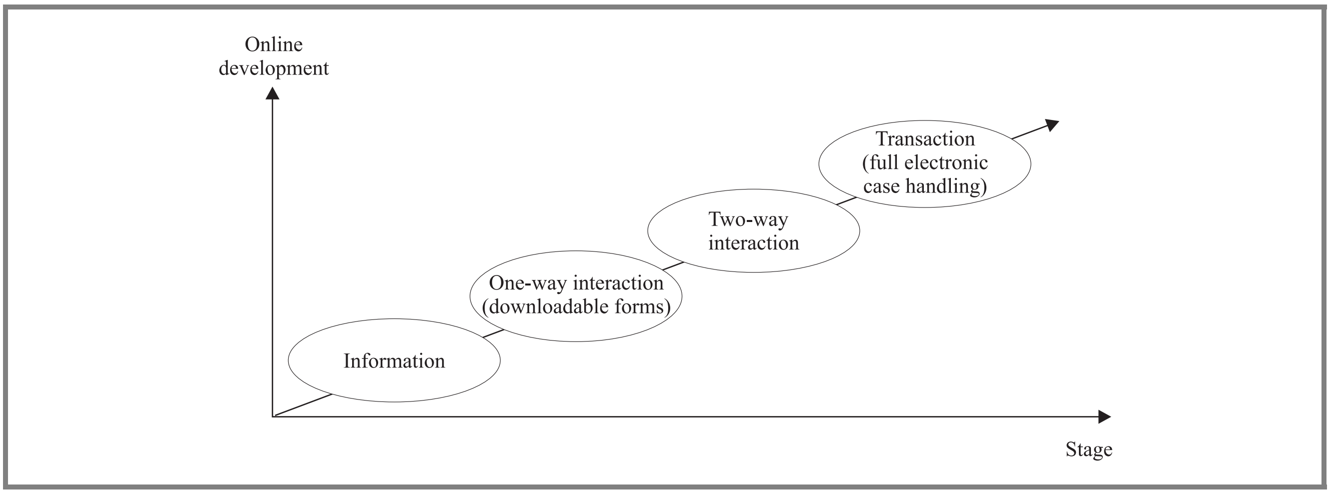


Fig. 1. Progress of online services for businesses and citizens.

the same today as they were yesterday [2]. Importantly, organizations and people can have different identities when working with different systems, or can have different identities when working with a single system, perhaps when working in different roles. While names and naming protocols are a critical element of identity, in that they give us the means to call out one identified entity from another, the underlying relevance, role, context and meaning attributed to a given named individual or company can only be gleaned by reference to other factors. The full measure of identity is a subtle and multi-faceted complexity because people and organizations exist in many social, economic, political, cultural and other dimensions all at once.

It is well understood that deploying an identity management solution has many dimensions and uncertainties. Nevertheless, governments are now faced with a complex set of challenges as they are asked to balance the need for security, privacy, citizen and business demands for online services, and the issuance of digital identities to make these services a reality. This is not a simple undertaking and must be supported by a complex framework of laws, policies, institutional decisions, business practices and ultimately, technology.

As articulated in the 2002 NECCC *Identity Management White Paper* [3], the vision of identity management for e-government solutions is to "support common identity needs of governmental and private transactions" and "reduce costs of government and enhance service quality". It is well understood that this vision must be achieved under an obligation to "preserve or improve individual privacy, name and identity related liberties and the security of identity information".

The issues involved in creating, using, changing and ending an identity involve technical, procedural, legal and policy dimensions. The advent of the information age has raised many of these issues anew. Current information management capabilities provide tremendous leverage in accessing, processing, manipulating and stealing information. This raises questions of privacy, security and fair information practices on the one hand, to be balanced against convenience of e-government service delivery, the need to identify and apprehend terrorists and fraud artists, and the need to interoperate across government and private systems on the other hand.

Every EU citizen and every EU company owns rights and obligations. Many of them would expect to be able to fulfill those rights and obligations wherever they are registered or work in the EU. Access to public services at pan-European level is a key aspect of this approach. Many research efforts in the last few years are devoted to the development of a mechanism for accessing those rights and meeting the obligations, which should be straightforward, easily understandable and accessible anywhere anytime within the member states.

Increased expectations of citizens and businesses, wanting to transact with their government through the interface of their choice, and the necessity of governmental entities to interact more effectively and efficiently with each other, have led to focusing for solution development on the capabilities available through the Internet.

2.2. Current state of the e-government in Europe

The improved delivery of public services is getting a very critical element of the wider economic strategy to modernize the EU economy. The European Commission's fifth annual survey of online government services in Europe [4], points out the impressive progress in developing and delivering public services online across the EU. The study reveals that over 90% of the public service providers now have an online presence, and 40% of basic public services are fully interactive. The service delivery gap between new member states and the pre-enlargement EU 15 is lower than many expected and could close very quickly. The availability and interactivity measures used in the survey show that EU's new member states have reached the level of the EU 15 from just two years ago.

The study goes further, analyzing the sophistication of the online public services provision, presented in Fig. 1, for the target user groups, citizens and businesses. The results reveal that the services for businesses reach an overall score of 77% for online sophistication and 58% are fully available online, while the services for citizens stay at the level of 57% for online sophistication and only 27% are fully available online. One of the main reasons for this significant difference is the fact that e-government services towards business are frequently revenue-generating services for governments. Additionally, the business processes, information systems and technical infrastructures are typically better developed than the ones used with citizens and therefore businesses easily adopt e-government services.

By developing an open identity management architecture to underpin e-government solutions, GUIDE will enable governments to offer higher quality services to businesses and citizens, thus reducing administrative costs and fighting the negative consequences of the virtual space, and will also contribute to an improved collaboration between different departments and the harmonization of e-government on a pan-European level.

2.3. Interoperability as enabler for European e-government development

The e-government has been developed so far in a very fragmented manner. E-government services are deployed by a multitude of public administrations at the national, regional and local level. Those services are islands of automation which cannot work together. This fragmentation may severely handicap the wide take-up and widest possible impact of e-government unless joining-up administrations and inter-linking online services is made possible through the interoperability e-government services.

According to the European interoperability framework [5], the term interoperability means "the ability of information and communication technology systems and of the business processes they support to exchange data and to enable the sharing of information and knowledge". Three aspects of interoperability have to be considered when setting-up services designed for more than one public administration.

Organizational interoperability concerns the definition of business goals, modeling of business processes and collaboration of administrations that wish to exchange information, but that may have a different internal organization and structure for their operations. The requirements of the users should also be addressed by making services available, findable, accessible and user-oriented. Semantic interoperability includes ensuring that the precise meaning of exchanged information is understandable by any other application not initially developed for this purpose and enabling systems to combine received information with other information resources and to process it in a meaningful manner. Technical interoperability concerns linking up computer systems and services, which includes key aspects such as open interfaces, interconnection services, data integration and middleware, data presentation and exchange, accessibility and security services.

One of the main objectives of GUIDE is to establish the EU as the global leader of e-government services through the enablement of an open architecture for identity management based on durable trans-national co-operation and consensus on a pan-European basis. It will be achieved by providing an architectural vision that integrates local, national, and international (pan-European) identity management services to establish a conceptual identity management grid, described below.

3. GUIDE open identity management architecture

3.1. State-of-the-art in the field of e-government and identity management

The research work carried out in GUIDE reveals that the different approaches to establishing e-government frameworks repeat some of the experiences of the enterprise domain, especially in that earlier versions focus on technical issues alone and later editions increasingly broaden the scope to organizational and policy issues. Early versions of e-government frameworks focused entirely on technical aspects, defining protocols and interfaces between systems. Over time and after several failures of architecture efforts the need for business driven approaches was widely accepted. This led to the development of layered frameworks deriving technical requirements from a pre-defined business strategy. Similarly the level of integration increases from rather low integration in early architectures to higher levels of integration in later editions.

Numerous solutions for e-government applications have evolved, each with their associated strengths and weaknesses. Most have focused only on offering a technical architecture, which neglects the incorporation of other aspects of identity management such as trusted third parties issuing and managing credentials, privacy, access control, risk and liability management. However, identity management is not just a technical issue. Identity management and perceived security are as much dependent on the context in which they are applied as on the architecture used. Moreover, every identity management solution to be implemented in the area of e-government faces the challenge to integrate smoothly with existing systems. The integration can be achieved through interoperability that can only be secured through the development of an open architecture.

Recent years have brought increasing research in the field of e-government and identity management. There are a growing number of projects in this research area, financially supported by the European Commission and uniting the efforts of leading European industrial and academic partners, e.g., the projects EMAYOR, HOPS, GUIDE.

GUIDE's mission is to lay the foundations of a generally accepted open identity management architecture for e-government on a European level. The research of GUIDE is focused on addressing the full range of technical, process, policy, legal and social issues that will allow this vision to proceed. GUIDE brings together the European industrial, financial and technical market leaders in e-government solutions, as well as leading academic institutes of the relevant scientific disciplines.

3.2. GUIDE architecture framework

Frameworks for e-government are in an early state of evolution. In this situation a look at the available solutions in the well developed enterprise domain allows adopting existing results and experiences in their development process. This is a disciplined approach to understand how components of an enterprise communicate, change, and function together as a whole [6].

An architecture is defined as a collection of independently useful systems that have been integrated together to achieve additional properties not associated necessarily with any of the individual systems. The strong focus is on communication and cooperation, and therefore the idea of interoperability between systems is paramount.

In the course of the development of the field, architectures have become more encompassing in that they do not only cover computer hardware and software, but increasingly as well organizational and business dimensions. They have as well become more sophisticated in their internal structure to respond to more demanding requirements especially for integration and flexibility.

GUIDE has assessed a number of industry approaches for architecture development such as Zachman and TOGAF8, which are implemented using Popkin system architect and RUP SE tools. At this stage of the project the Zachman model is adopted as a general framework for developing a methodology, specific to the requirements of an open identity management architecture, as it provides a starting point and an industry standard approach. However, GUIDE's research is not confined to that model. During the research and development process other approaches will also be used if considered useful and effective.

The Zachman enterprise architecture framework [7] is shown in Fig. 2. Each cell represents the intersection of a particular focus and a perspective. Each focus (the question what, how, where, who, when, and why) is depicted in a column and each perspective (point of view) – in a row. The perspectives define the point of view or the level of abstraction for the information contained in the cells. The information and models within a single row represent a complete description of the architecture from that perspective. Each column captures all of the architecture knowledge for the particular question being asked, the focus. The total architecture knowledge for each focus is obtained by isolating each focus and defining the artefacts for each perspective within it.

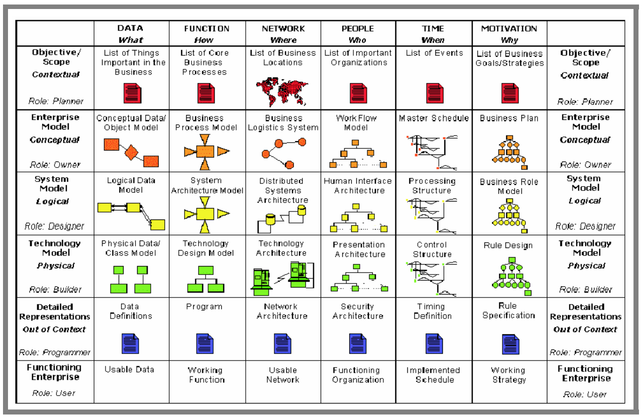


Fig. 2. The Zachman enterprize architecture framework (*copyright: John A. Zachman, Zachman International*).

Service oriented architectures (SOAs) are state of the art in enterprise architectures. Conceptually, SOA represent a model of loosely-coupled applications working together by exposing services to each other. Business wise, services are expressing data- and function-services that one party can offer other parties to use. Technologically, SOA consists of a group of emerging standards that defines protocols and creates a loosely-coupled framework for programmed communication between different systems. Web services are a specific implementation of a SOA, i.e., a method which enables an application to be invoked by other applications by receiving and sending data in standardized XML.

Taking into consideration the above mentioned, the essence of the GUIDE architecture is foreseen as a service oriented architecture, given the obvious requirements for "loosely-coupled" systems, independence of implementation and location, etc. Furthermore, the only real candidate implementation of SOA currently is the web services model, and this is envisaged as the most likely physical perspective candidate for the GUIDE architecture. Given that there will be a need for a highly secure approach, it is envisaged that the service oriented security model (SOSA) as currently delivered by the emerging web services security model (WSSM), will be overlaid onto the basic WS architecture.

The design and development of the GUIDE open identity management architecture is driven by eight key political and functional axioms, to which further research is designed to add more knowledge and insight. These axioms are as follows:

1. European open identity architecture: "A European open identity architecture will be defined".
2. External applications: "All identity data is produced and consumed through applications outside the identity grid".
3. External data: "A significant amount of identity data will always stay outside the identity grid".
4. External transactions: "A significant amount of identity transactions will always be done outside the identity grid".
5. Data ownership: "Each functional element of identity data within the identity grid will have clear data ownership and data obligations".
6. Identity services: "Applications outside the grid will interact with a set of *attribute service providers* within the identity grid".

7. EU governance: "The architecture will conform to the overall EU regulatory and legal framework and system of governance".
8. State governance: "Each member state will have governance over attribute services operating within their boundaries, and the identity data underpinning these identity services".

As the development of the architecture progresses these will be supplemented, at a more detailed level, by a set of "architectural principles", including:

- Guide will adopt service oriented security architecture as its underlying architecture.
- Guide will align with and complement other EU information society initiatives, such as IDA.
- Guide will align with and complement emerging industry initiatives and standards in federated identity management, such as the liberty alliance.
- Guide will observe the principle of subsidiarity.
- Guide will be inclusive of all member states identity management requirements where these do not conflict with the majority position or can be provided as an optional feature.
- Guide will develop a proportional response to the overall requirement for pan-EU identity management, with the aim of providing maximum benefit for minimized effort.

GUIDE's strategic vision is to develop an architecture that integrates local, regional, national, and pan-European identity management services in an interoperable manner that allows accommodating the requirements of member states. As such the GUIDE architecture is consistent with the principle of subsidiarity. It is based on a federated information infrastructure model that respects the sovereignty of member states in identity management issues, rather than a hierarchical one.

In this relation GUIDE is conceived as providing a pan-European federation of identity federations, where the architecture of each constituent federation is deferred to the "owners" of the federations, including member state governments and commercial organizations. GUIDE focuses on how these federations should interoperate, such that in totality the whole can be conceived as an identity grid or identity network for Europe.

The GUIDE identity management grid presented in Fig. 3 is a high level visual representation of the concept for pan-European architecture, derived from the axioms defined above.

Central to the model is the identity holder. It represents the holder being in control of the identity. This is one of the key principles of GUIDE. The identity holder can be a physical person such as a citizen, but may also be a legal entity such as an organization, or automated agents such

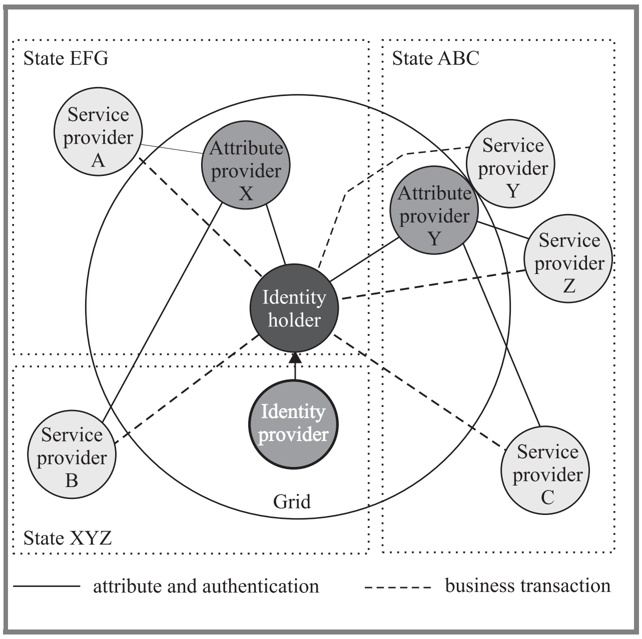


Fig. 3. Identity management grid conceptual model.

as web services. It should be noted in this graph, that the identity holder is not part of the grid. However, since the user is "in control" of the identity, it should be represented in the middle. One could view the identity holder as standing "above" the grid.

The identity grid contains identity data elements. They are released under a specific set of circumstances, governed by a number of protocols. The grid is represented by the large circle and all elements that lie within it. The identity grid will be a set of physical services and data elements, controlled by procedures and policies (Axiom 1).

The service providers and their applications sit outside the grid. Service provider applications connect into the grid in order to request and process identity data elements (Axiom 2). In addition to e-government applications, commercial applications may ultimately also be allowed to utilize the grid, obviously when satisfying the GUIDE principles.

Some identity data elements will remain outside the grid for a certain period (Axiom 3). Processing of this data will be governed by protocols derived from principles on data protection, security, retention and so on. Data elements exist in databases alongside each application. Furthermore, many (non-e-government) applications and databases are outside this visual model.

The model focuses on e-government applications and their use of identity data. Hence a number of identity transactions (such as commercial transactions, and e-government service providers that are not connected to the grid) will remain outside the grid (Axiom 4).

One of the most critical activities in identity management projects is the establishment of protocols on processing of data elements. This is essential for ensuring of accuracy and quality of data, as well as for meeting other social, legal and regulatory requirements. Data subjects will always be in control of their data – the citizens can declare how their data may be used and its accuracy maintained in certain cases. In addition, there will be legal entities that are accountable for legitimate data management (Axiom 5).

There will be a set of identity services within the grid (Axiom 6), in the form of identity service providers. Each identity service provider will be governed by an individual state and "certify" the identity of the identity holder. There may be multiple identity service providers, just as there are multiple unique IDs in various countries. In essence, each country of the EU represents a single identity service provider. Furthermore, the arrows represent the conceptual flow of data to and from identity services within the grid. Key to GUIDE is that identity service providers do not interact with service providers – an attribute service provider is set up as a mediator – with the purpose to keep the identity holder in control as much as possible.

Attribute service providers deliver the required identity attributes credentials and authentication services to the application service providers. This will depend on the strength of authentication that is required for a specific application. There may be more attribute service providers that an identity holder is affiliated with, and a service provider and attribute service provider may originate from the same entity (service and attribute service provider Y in Fig. 3).

Within the identity grid itself, the overall governance of the grid will be driven by regulatory bodies that follow rules derived from EU legislation (Axiom 7). Within the overall EU governance, individual states will carry governance over the identity data and services within their boundaries (Axiom 8). The pan-European aspect of different states is represented in the model by the dotted boxes.

3.3. Innovative interdisciplinary approach to identity management for European e-government development

The formation of identity in all its dimensions is conditioned by the institutional, policy, legal, and regulatory frameworks in (or against) which it evolves. Information and communication technology broadens the scope of interaction of these frameworks, which means that institutional, political and legal frameworks that function as the drivers and sources of identity interact on a global scale. In this context identity drivers diffuse across institutional environments, borrow from each other and create amalgams that correspond to common patterns of identity formation, while adapting to the specific social environments within which they operate. This is to say that the formation of identity is a process that is mediated by antecedent, yet evolving, institutional, political, legal, and regulatory forms. This mediation is of fundamental importance in the emergence, or lack thereof, of factors that encourage the uptake of e-government services. A critical question then concerns analysis of the evolving topology of the institutional, political, legal, and regulatory sources of identity formation. Such a question can only be answered by approaches that remain sensitive to the dialectics of integration and reproduction of difference that mark contemporary EU history.

Much of the research work within GUIDE focuses on institutional, policy, legal and sociological frameworks underpinning identity management in order to identify conditions for, and obstacles to, EU-wide take-up of e-government services [8]. Central to the research is an understanding of critical organizational and political aspects of identity management. An analysis of the legislative landscape at both, national and EU level, is also undertaken, as it will give the "enabling legislative framework" that will shape the paths of development of the open identity management architecture of GUIDE. The socio-economic, ethical and cultural differences that drive identity formation will be studied and identified as well.

GUIDE, being an integrated project, stresses on "integration" of the conceptual and research components of the project. All documents, produced as a result of the in-depth studies of the institutional, political, legal, socio-economic and policy aspects of identity management, are continuously analyzed and requirements towards the developing open identity management architecture are being formulated. These requirements are directed to the identified architecture pillars and their aspects, including identity data (security, confidentiality, integrity, availability, privacy, intra- and inter-state identity data transactions, data holders, data users); identity management services (security, accessibility, user interface, standards, protocols, service providers, service users), identity management processes (security, standards, protocols, process providers, process users), interoperability, multilingualism, etc. Each requirement is given priority expressed in at least three basic ratings: compulsory, important and nice-to-have. The prioritization is performed by both, the industrial experts responsible for the architecture design and the academic researchers studying the settings, in which it will function. The importance of each requirement for the proper functioning of the architecture is primarily considered. The assigned priority is dynamic and might change with the progress of the GUIDE research and architecture design.

The key research findings and their implication for EU and member states will be synthesized and presented in one of the project deliverables titled *GUIDE Policy White Paper on Identity Management*.

3.4. The verification process

GUIDE open identity management architecture will consist of a collection of identity management services integrated by a combination of technical and non-technical compliance criteria. The verification of this integrated architecture will be realized by performing different trial tests. A number of innovative identity management services will be developed and demonstrated in parallel with already existing identity management services, as part of an overall open identity management architecture. The identity management services will be trialed with a variety of e-government applications to demonstrate the effective meeting of user requirements for these services, as well as across different geographic scenarios to ensure the broad capture of requirements and demonstrate full effectiveness in a wider variety of settings. GUIDE trials will not only provide important verification of and give inputs for the open identity management architecture being developed. They will also provide tangible evidence that the open architecture can be implemented and will provide a valuable reference case in discussions with governments and other interested parties.

4. The future of identity management in the e-government perspective

There are different points of view, often quite contradictory, concerning the principles which should guide the policy, legal, business and technical architectures for identity management systems and practices. However, it is necessary to devise innovative methods and approaches that support a balanced reflection of each of the competing interests.

The involved decision makers have to carefully consider the policy, technical, legal and business ramifications of identity management at all levels – local, national, European. The technical architectures chosen are not policy-neutral, in that they carry with them certain explicit or implied assumptions about the roles and expectations of users. In addition, the policy and legal approaches are charged with potential for missteps and controversy. However, it is clear that the basic drivers toward implementation of better identity management systems and methods will move European states and other stakeholders toward creating more, bigger and broader systems.

5. Conclusions

GUIDE's overall goal is to create the main critical requirements and principles for open identity management architecture development that will support EU e-government services interrelations and interoperability, based on durable trans-national co-operation and consensus on a pan-European basis. The vision to create a pan-European service-oriented architecture will allow the dynamic interoperable e-government services and applications throughout Europe, whilst preserving state subsidiarity. The innovation in GUIDE is the attempt to research and define how existing identity management services can inter-operate with new identity management services in a pan-European setting, and the adopted encompassing interdisciplinary approach to identity management, seeking to overcome the existing fragmentation of identity management initiatives.

Through achieving the scientific, technological and socio-economic goals GUIDE will contribute towards initiatives that will ultimately deliver multiple benefits to governments, citizens and business. Identity management can be applied to many different e-government services solutions by creating a consensus on European identity management architecture. Identity management services can be turned into key contributions to the further advancement of e-government throughout Europe to create the European market leadership.

Acknowledgements

This work has been performed in the framework of the IST project GUIDE (IST-2003-507498), which is funded in part by the EC. The Authors would like to acknowledge the contributions of their colleagues from BT Limited, Siemens Schweiz AG, Visa International Service Association, Center for Technology and Innovation Management, Crealogix AG, ELCA Informatique SA, Budapest University of Economic Sciences and Public Administration, PricewaterhouseCoopers BV, Cyota, DeCon APS, Eesti-Taani Kommunikatsioon Ltd, Infonic Ltd, Modirum Oy, NetSmart SA, Tecnologia E Ingeneria De Sistemas y Servicios Avanzados De Telecomumicacion, Eidgenoessische Technische Hochschule Zuerich, The University of Surrey, European Institute of Interdisciplinary Research, Research Center of the Athens University of Economics and Business, ERASMUS Universiteit Rotterdam/Rotterdam School of Management, The Chancellor, Masters and Scholars of the University of Cambridge, Sofia University "St Kliment Ohridski" and DL Legal. The authors would like to acknowledge that they are solely responsible for this document and that it does not represent the opinion of the European Commission, and that the Commission is not responsible for any use that might be made of data appearing therein.

References

[1] IDABC eGovernment Observatory, "The impact of e-government on competitiveness, growth and jobs", European Communities, 2005, http://europa.eu.int/idabc/egovo

[2] The Open Group, "Identity Management", a "White Paper", 2004, http://www.opengroup.org

[3] The National Electronic Commerce Coordinating Council, "Identity Management White Paper", in *NECCC Ann. Conf.*, New York, USA, 2002, http://www.ec3.org/Downloads/2002/id_management.pdf

[4] European Commission Directorate General for Information and Media, Capgemini, "Online availability of public services: how is Europe progressing?", 2004, http://europa.eu.int/information_society/soccul/egov/egov_benchmarking_2005.pdf

[5] IDA, "European interoperability framework for pan-European eGovernment services", Brussels, 2004, http://europa.eu.int/idabc

[6] K. Stefanova, D. Kabakchieva, and L. Borthwick, "GUIDE open identity management architecture design – key contribution to the further advancement of egovernment throughout Europe", in *5th Eur. Conf. eGovernment (ECEG 2005)*, Antwerp, Belgium, 2005, pp. 377–386.

[7] J. A. Zachman, "A framework for information systems architecture", *IBM Syst. J.*, vol. 26, no. 3, pp. 276–292, 1987.

[8] GUIDE Consortium, "GUIDE Description of Work", Ver. 4.5, Oct. 2004.

Kamelia Stefanova is a Project Manager at the Centre for Information Society Technologies, Sofia University. Since 1984 she has been a lecturer in information and communication technologies, teaching management information systems design. Doctor Stefanova has developed great expertise in the fields of e-government, banking and finance, knowledge economy, participating as an active researcher in many European projects.

e-mail: kamelia@fmi.uni-sofia.bg
Centre for Information Society Technologies
Sofia University
125 Tzarigradsko Shosse Blvd.
1113 Sofia, Bulgaria

Dorina Kabakchieva is a Project Manager at the Centre for Information Society Technologies, Sofia University. She has been actively involved in European projects within the e-government and e-business scientific areas. She has the M.Sc. degree in educational and training systems design (Twente University, the Netherlands) and the M.Sc. degree in electronics (Technical University – Sofia). Currently she is a Ph.D. student in business information systems.

e-mail: dorina@fmi.uni-sofia.bg
Centre for Information Society Technologies
Sofia University
125 Tzarigradsko Shosse Blvd.
1113 Sofia, Bulgaria

Lia Borthwick received the B.A. (hons) degrees in English literature and the M.Sc. in telecommunications business. She is an e-government specialist with several years experience of working in e-government and research projects. She is currently the Programme Director for Europe's largest research project on identity management for e-government. The project involves 23 organizations from industry and academia across 13 countries. She has presented at conferences in Europe, USA and India on e-government, intranets, e-business and knowledge management as well as a speech in the House of Lords to senior members of the UK government and has given a keynote speech at the World Bank's knowledge for development conference with government representatives from China, India and Brazil. Since 1997, she has had several articles published in UK trade media and has also been interviewed extensively on the issue of citizen identity management in a variety of quality and trade publications throughout Europe.

e-mail: lia.borthwick@bt.com
BT Ltd., BT Centre
81 Newgate st
EC1A 7AJ London, UK

Paper

A way of integrating deaf, hearing- and speech-impaired people into modern communication society

Marlene Hilzensauer

Abstract— This article deals with the exclusion of deaf, hearing- and speech-impaired people from our modern communication society, as they are unable to use the phone. This situation leads to discrimination and disadvantages in their everyday lives. One solution to this problem is the implementation of so-called relay centers which act as go-betweens; with the aid of communications assistants and interpreters, a direct conversation with a hearing person becomes possible. The Center for Sign Language and Deaf Communication has developed a concept of such a relay center for Austria. The main innovations are the integrative technological approach (trying to integrate all possible communication devices) and the embedding into an Internet portal, including additional services for the target groups. An overview of the activities carried out during the 6-month preparatory project is given, and the results are described.

Keywords— telecommunications, communication society, integration, inclusion, deaf, hearing-impaired, speech-impaired.

1. Introduction

In our modern communication society, most people prefer to use the phone for long-distance communication. Whether at work or at home, if you have a question, if you need some information, want to order some product or just simply chat with someone else, you will usually reach for your phone – even more so with the introduction of the mobile phone which allows people to be available around the clock.

However, probably because this way of communication is such an integral part of our lives, we hearing tend to forget that there exist several groups of people who are unable to communicate via the phone: the deaf and the severely hearing-impaired because of their acoustic inability, the speech-impaired because they have trouble with their articulation. They can either manage only with difficulty and frequent misunderstandings, or they are even barred completely from using the phone. The latter group is forced to depend on friends, relatives and colleagues to make their phone calls for them. In Austria, for example, this concerns more than 100 000 people.

There are negative consequences not only for their private lives, but especially for their professional lives. Companies are wary of hiring employees who cannot make phone calls, e.g., to customers and suppliers, so members of the target groups may be turned down by prospective employers; with deaf people, there is an added fear of difficult face-to-face communication.

As access to telecommunications may be regarded as a civil right, some countries have taken measures to fight this discrimination: one possible solution is the implementation of relay centers.

2. How a relay center works

2.1. Basic functions: text and video

If deaf, hearing- or speech-impaired people are unable to use the phone without assistance, a relay center is needed as a go-between. Basically, there are two types of relay center: text-based or video-based.

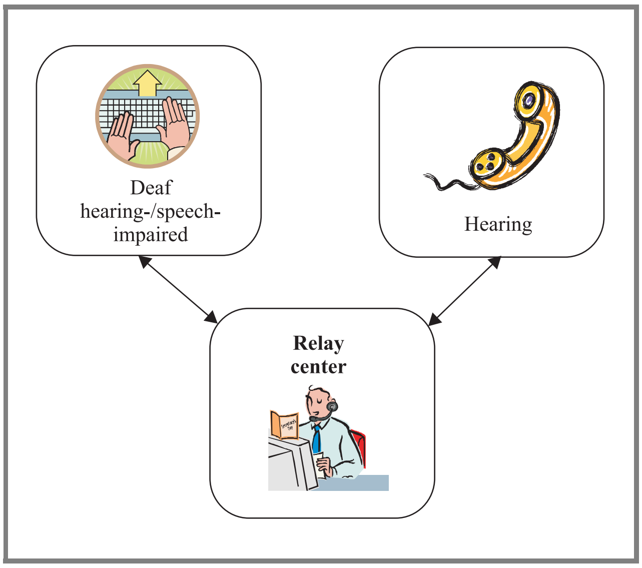


Fig. 1. How a text relay service operates.

Let us assume that a deaf person wants to contact a hearing colleague: first, s/he has to establish communications with the relay center. If s/he is using text, s/he will probably either do this via a computer or a textphone[1]. The communications assistant in the relay center will then phone the respective hearing person and read to them what the deaf person has written. The spoken answer is written down and sent back to the deaf person (Fig. 1). Naturally, the com-

[1] A text telephone (TTY) – also called telecommunications device for the deaf (TDD) – is a special phone for deaf people, where text is transmitted via the normal phone network. Two people using textphones can communicate in real-time, as every character that is typed shows up instantly on the other person's display. Nowadays, no special device is needed as modern computers can emulate a textphone. For more information about TTYs [1, 2].

munications assistants must be well-trained (for example, they are required to type at least 60 words/minute in the USA) and they are obliged to keep every conversation totally confidential (offenders are immediately fired). This kind of relay center is called text relay service (TRS).
With video, the same principle applies, but the user sits in front of a camera, e.g., a web cam or a videophone (Fig. 2).

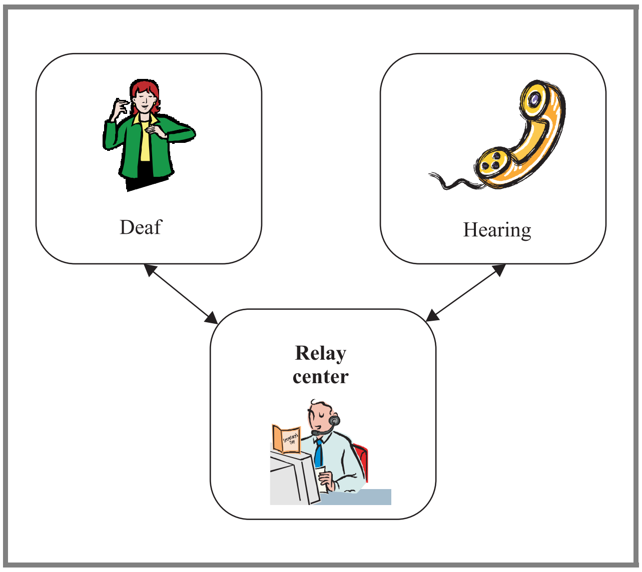


Fig. 2. How a video relay service operates.

This is especially interesting for deaf people, as with modern broadband technology, they can use their native language, a sign language. The deaf person signs their part of the conversation which is translated by an interpreter in the relay center. Again, the interpreter makes a phone call to the hearing party the deaf person wants to talk to, listens to the answer and signs it back to the deaf person. The term for sign language video communication is video relay service (VRS).

2.2. Special functions

Voice carry over (VCO). Some people with a hearing-impairment – especially those who are late-deafened and normally do not sign – may prefer to speak themselves instead of typing their part of the conversation. In this case, they will only receive the answer from their hearing conversation partner in written form.

Hearing carry over (HCO). This feature will probably be used mostly by speech-impaired people. While they cannot speak intelligibly, and therefore have to resort to writing, they can listen to the spoken answer themselves without any problem.

The CapTel phone. Basically, this is a variation on HCO. An American company developed a special captioned telephone[2] for hearing-impaired users, which features an in-built display. It can be used as a normal phone or – if the user so chooses – in connection with a text relay service. The target group are the hard-of-hearing. Many of them can make phone calls themselves, but with varying success, depending on factors like environmental noise or whether they know the person they are listening to. With CapTel, if the text relay feature is activated, everything the hearing person says is repeated by a communications assistant and typed by speech recognition software (the communications assistant is necessary because speech recognition software works much better if it has been specially trained to recognize a certain person's voice, thus making fewer mistakes). In this way, the user can check with the written version if s/he does not understand well or if anything is unclear.

Speech-to-speech. This is another feature for speech-impaired users which is offered by some relay centers, e.g., in the USA. A specially trained communications assistant listens to what the user articulates and repeats everything s/he says to the conversation partner at the other end of the line.

Remote interpreting. Also known as distance interpreting, this is a special, expanded version of video relay for deaf people. Alternative terms are video relay interpreting (VRI) or videophone interpreting (VPI). Although a physical interpreter is always preferable, sometimes this is impossible (one of the main reasons is the shortage of qualified sign language interpreters in many countries). An alternative is to use the interpreter in the relay center: this works similar to video relay, but the deaf and hearing conversation partners are sitting in front of the camera together, while the interpreter translates what is said/signed via video connection (Fig. 3).

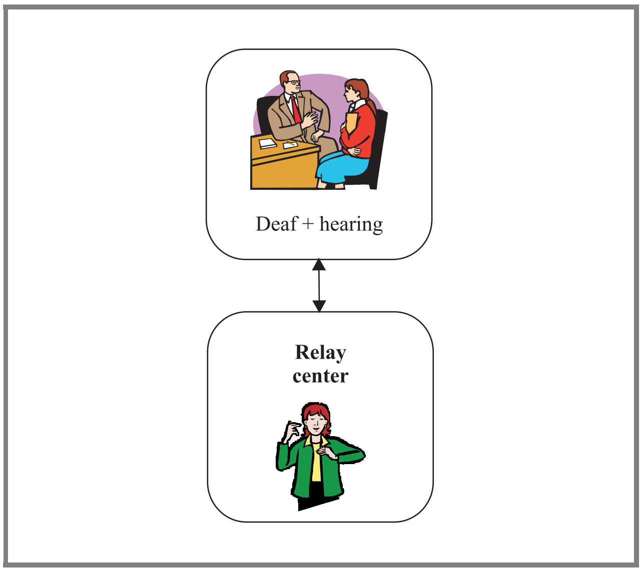


Fig. 3. Remote interpreting.

Communication Service for the Deaf (CSD) provides such a service called online interpreting[3] and calls it a *quick way to get a qualified interpreter in place for an effective*

[2]Cf. http://www.captionedtelephone.com/index.phtml

[3]Cf. http://www.csdinterpretingonline.com/index.html. On this homepage, you can also find a demo video of remote interpreting.

communication with a deaf or hard-of-hearing individual who may require use of sign language [3].

Remote interpreting may not be suitable for every interpreting situation; for instance, in court or at the doctor's, an interpreter should be present, because any misunderstanding could lead to grave consequences. Basically, remote interpreting is ideal for brief conversations where booking an interpreter would be both impractical and too expensive. The German video relay service "Telesign" names inquiries, appointments, brief meetings and instructions as ideal for remote interpreting, but warns of using this service for longer meetings where several persons are present, e.g., psychically difficult conversations or company meetings [4].

An overview of which situations are regarded as difficult by an experienced interpreter is listed in [5]:

1. *Meetings/dialogues with more than 4 persons present in the room, if the videophone is without an extra microphone or other equipment upgrading the sound.*
2. *VPI from a location with a lot of back ground noise.*
3. *If the users of the videophone have to be mobile, more around in the area, stand up, sit down, etc.*
4. *From classrooms or courses where the blackboard or AV-equipment is used.*
5. *Outdoors.*
6. *When the situation [requires] that the sign language interpreter can see all the participants or for other reasons where the sign language interpreter needs to get the visual information at the spot.*

For the interpreters themselves, video relay and remote interpreting demand high standards of them: not only do they have to adapt to a multitude of different interpreting situations from one call to the next, but there are even some changes in the sign language used for interpreting, to mirror the special situation: for instance, Danish interpreters have altered the sign for "I" by no longer pointing to themselves but rather to the microphone of their headset, thus emphasizing that the hearing conversation partner has said something [6].

3. Advantages of a relay service

To people who can use the phone themselves without a second thought, relay services may seem a slow and awkward way of communicating. What is important, though, is that it makes direct communication possible. One need not wait for a person's answer, but can interrupt them, ask a question or clarify at once everything that one may not have understood.

While this system is certainly not perfect, relay services nevertheless help the target groups to regain their independence with regard to telecommunications, and allow for a better inclusion into society, while at the same time improving their job chances. This is especially true of video relay for deaf people.

When talking with hearing people, the question if relay centers are really necessary keeps coming up. Usually, written communication in the form of SMS, e-mail, fax, etc., is regarded as an easily accessible and much cheaper alternative for relay services. These forms of electronic communication are an important part of our lives, but they are no substitute for a relay service. First of all, you need the respective number or address, which may not always be available. Then there is the problem of confidentiality: with a fax, for example, you never know who might have read it besides the recipient. There is also no proof that it has reached the intended person – SMS and e-mails have been known to arrive days later or even not at all. Sometimes, such a breakdown in communication can have serious consequences: take, for instance, deaf parents who want to contact a doctor because their child is sick. They send a fax to the doctor's office and wait patiently for an answer, while the doctor is on holiday – simply because they cannot listen to the respective tape message. With SMS, there is also the size limit of 160 characters, which makes a longer communication rather difficult.

Another problem is the hearing communication partner – imagine, for example, discussing some question about your income tax with the tax office via SMS. Probably the hearing party will answer once or twice, but will not write back ten times or more. Even if they do, the information will usually be condensed or not complete, because people do not want to write so much.

The need for sign language. For the deaf, there is still another barrier: most hearing people are not aware of deaf people's difficulties with written language. According to their way of thinking, when somebody cannot hear, they can always read the same information; and if they cannot speak, they can still write their message down. However, mostly due to education methods which are not tailored to their unique needs, many of the deaf have trouble understanding longer or more complex texts. In writing, they make a lot of mistakes, and because they know this, they do not like to write to hearing people. What the hearing forget, is that any written language is usually a second language for the deaf. Therefore, the possibility of using their own language, i.e., a sign language, comes as a relief to them.

A second reason for video relay is the ease of communication. Signing is much faster than writing, especially if the user in question does not type very well.

The advantages of video communication are not limited to deaf people, however. Hard-of-hearing people may profit from lip-reading, while showing some object or watching the other person's emotional reaction is possible for all people.

Advantages of remote interpreting. Although there are some drawbacks – for example, the quality of the video connection may be less than perfect – remote interpreting is ideal for rural or remote settings (e.g., in Norway or

Australia, where the next interpreter is far away), and spontaneous questions or conversations. Booking an interpreter for an on-the-spot meeting is an impossibility (especially if there are only a few interpreters available, you need to book them far in advance), and often the expense is too high for such a short duration. On the interpreters' side, a lot of time is saved because they do not have to drive to a certain location – this takes often longer than the whole interpreting session. During the same period of time, they can deal with multiple requests.

A special application is the use of video mobile phones: although the video quality tends to be worse than with some stationary device, the mobility is a bonus. As the Swedish company Netwise shows in one of their demo videos, you practically have an interpreter in your pocket[4]. In this video, a deaf lady was able to make her own appointment at the dentist's and to communicate all her wishes by simply putting her mobile on the receptionist's desk and signing into it, while the receptionist listened to the interpreter's voice.

In Australia, it is regarded as discrimination, if deaf people are denied access to an interpreter; one way of guaranteeing this is the use of remote interpreting:

Failing to provide a Deaf person with access to an interpreter is unacceptable in the era of Disability Discrimination legislation. Such legislation underlines obligations regarding the provision of communication access for all people. Employers, colleagues and others who work with Deaf people must not accept that inadequate numbers of interpreters can dictate access to communication for and with a Deaf person. Alternative methods of providing communications access must be found. VRI is one possible option to help improve access. With careful planning, marketing, close work with interpreters and consumers, VRI could become an invaluable service, if not a "life-line" for many people [7].

4. The preparatory project for an Austrian relay center

The idea of implementing an Austrian relay center dates back to a workshop hosted by the Center for Sign Language and Deaf Communication (Zentrum für Gebärdensprache und Hörbehindertenkommunikation, ZGH) in 2000. During the workshop which dealt with preparing a European deaf network for information and communication (for a final report [8]), some of the participants related their experiences with relay centers in other countries, e.g., in Sweden and Switzerland. Encouraged by their example, the Center for Sign Language and Deaf Communication decided to found a similar business in Austria. In the fall of 2003, we started searching for the necessary funding; finally, the Carinthian branch of the Social Welfare Office agreed to fund a three-month preparatory project, starting with August 2004 (this project was later extended for another three months because of the scope of the project and the work involved). Further assistance came from the university initiative "Build!", which helps graduates to found their own businesses and provides information and support. The main objectives of the preparatory project were the following:

- First, we were to collect as much information as possible on existing relay services worldwide and find out about "best practice".
- Second, we wanted to have a look at possible technical solutions.
- Third, we had to provide information to the Austrian deaf community and to the interpreters, and to initiate talks with the government.
- Fourth, we were to look at the legal situation in Austria, whether some law could provide a basis for a relay center.
- Fifth, we had to identify possible partners (both for the technical/strategic and the operative/contents side of the project).
- Sixth, we had to develop a concept for a relay center for Austria.

4.1. International relay centers

As for the European countries, the contacts that had been established during the workshop in 2000 proved very useful. Both Switzerland and Sweden did not only offer a wealth of information, but also invited the project team to their respective countries to have a look at their centers (a text relay in Switzerland, a text relay and a video relay in Sweden) and their technical solutions. With the USA, an Internet search led to lots of e-mail contacts and finally to personal meetings with representatives of US-based relay services (Sprint/CSD[5], Hamilton Relay[6], and Hands On Video Relay Service[7]). The head of the Center for Sign Language and Deaf Communication had the chance to visit a relay center run by Sprint/CSD in the USA. We tried to get a clear picture of the organization and to evaluate the procedures used in the relay centers whether they could be included in or adapted for the Austrian concept.

The Internet provided even more information on international relay centers, from diverse countries like Australia, New Zealand, Spain, etc. Very useful were reports about experiences with existing services, e.g., [9, 10], as well as publications like [11] to get an overview of what can be done.

[4]Netwise (formerly Envilogg) provides, e.g., multimedia technology which can be used for relay centers. Cf. http://www.netwisecorp.com/default.aspx?id=3

[5]Cf. http://www.sprint.com

[6]Cf. http://www.hamiltonrelay.net/

[7]Cf. http://secure.hovrs.com/VRS_SSL/hovrs.aspx

We also sent out a questionnaire to all the relay services that could be found; although some of the information we asked about was regarded as confidential, we did receive answers which proved very informative and were included in an overview table.
All the information has been compiled in a final report for the preparatory project [12].

4.2. Technical solutions

We soon realized that Austria's delay in implementing a relay center also had some advantages: in other countries, text and video relay were usually provided by separate centers. Text relay had come first, and when video finally appeared, often new companies took over and offered this service. Another problem was that even when new technology was used, there had to be some provision for users who still had the old equipment; for instance, outdated brands of videophones had to be included, if these users were not to be left behind.
Based on all the information we had gathered, we wanted a single product which would include video, text and speech. Those features could be used in any combination, so that the users were free to choose whatever suited their individual needs best. With text, it was important that every character was transmitted as soon as it was typed, similar to a textphone (not as it is usually done when chatting to someone, where you have to wait until they have typed a whole sentence or even paragraph).
Looking at the various technical solutions which were used, the software differed a lot, also with regard to costs and maintenance. In the course of the project, we made a first selection, based on our experience with best practice and what was needed for our Relay Center Austria. The field was narrowed down to six possible choices. These were to be tested during a field trial, together with the target groups and the sign language interpreters.
As the funding for the trial period took longer than we had expected, some of the companies in question generously let us try out their products in a small field test carried out by the Center for Sign Language and Deaf Communication. Special emphasis was given to qualities like easy installation and user-friendly handling, costs for the users, and reliability. Another factor was the geographical location of the companies involved – the closer to Austria, the better; also, with German-speaking companies, manuals and other materials did not need to be translated.
As all of the solutions had advantages and disadvantages (for a discussion of the products involved and our test results [12]), we decided to take an existing software and add some modifications to fulfill all our needs. If possible, it should be computer-based and web-based. The latter is important, because end users should need only a minimal installation, ideally none at all. The software will be integrated into an Internet portal, which provides additional services (cf., Subsection 4.6).
It is important, that the communications assistants or interpreters can take calls from different devices at a single workstation – although it might be possible to make do with a computer plus textphone, fax, etc., this would only serve to complicate an already demanding work.
As for contacting the relay center, the majority of the users will probably use some kind of computer. However, as we do not want to exclude anyone, the users should be able to use any device, even older ones like videophones, textphones, etc. Of course, computer-based solutions will be implemented first; older or rarer devices will be added later on, probably in separate smaller projects, in the order of their priority. An alternative for older people who do not want to use a computer could be access via a videophone which is connected to the TV-set.
What is important is that people can contact the relay center not only from a stationary device, but also from a mobile one (e.g., a mobile phone).
The technical base of the relay service should also offer an automatic conversion of different text formats; for instance, changing a fax message into an e-mail (because the deaf often use devices like a textphone or fax which are not owned by all hearing people).
Additional functions are, e.g., changing the layout of the software according to personal preferences or needs (larger print for partially sighted people), the possibility to save text communications and maybe even video calls, some visual alert signaling incoming calls, and an inbuilt answering machine.
Provision must also be made for users who need special devices, ranging from, e.g., deafblind users with a Braille keyboard to the hard-of-hearing who use induction loops.

4.3. Information of the target groups and negotiations with the government

The Center for Sign Language and Deaf Communication had naturally had contacts to the deaf community before, but they were intensified for this project. The project team traveled all over Austria to present the concept of the Relay Center Austria to the deaf and hard-of-hearing; information was also included in deaf newsletters and sent to the local deaf associations in the form of sign language videos. The Austrian Deaf Association (Österreichischer Gehörlosenbund, ÖGLB) which had demanded a relay center for years, helped to disseminate the information as well. Some of the deaf in Klagenfurt and Vienna also had the chance to test some of the technical solutions. The deaf's feedback on the concept of the Relay Center Austria as well as the technology was very important for us. The same holds for the sign language interpreters, who were informed of our plans and asked for their cooperation as well as for some input on working conditions for video interpreting (both relay and remote interpreting).
As for the government, we presented the project to local politicians (e.g., the Landeshauptmann, the head of the province of Carinthia) and to the Austrian government (mostly, the Federal Ministry of Social Security, Generations and Consumer Protection and the Federal Ministry of

Transport, Innovation and Technology). One of the problems is that in Austria, responsibility for people with special needs is split between several authorities: only vocational matters are dealt with by the Social Welfare Office. As we did not want to offer a relay service only for vocational issues, some cooperation between the different authorities is necessary. Because it is a technological project, the Ministry of Transport, Innovation and Technology is involved as well.

In February 2005, at the end of the preparatory project, a presentation was held at the University of Klagenfurt. All the Austrian deaf and the hard-of-hearing, as well as local and national politicians, were invited. After the modus operandi of a relay center as well as our ideas for an Austrian version had been described, video connections to the neighboring province of Styria, to Switzerland and to a relay service in the USA were established to demonstrate different technical solutions. Thus, the audience could get an impression of the video quality and see a relay service in action.

4.4. The legal situation

Internationally, relay services are usually funded on some legal basis, because the expenses are too high to be negotiated anew every few years (relay centers that were run on a project basis normally were discontinued when the project ended). An overview of international funding – both in Europe and the USA – can be found in [11, 13, 14]; the situation in the Scandinavian countries is described in [15].

Most countries choose one of two alternatives. The law in question is either a disability or antidiscrimination law – e.g., the Americans with Disabilities Act (ADA) in the USA – or the right to equal access to telecommunications, e.g., in Switzerland or in Germany. Because we had been made aware by the Swiss of similarities between the Swiss and the Austrian telecommunications laws, we engaged the Vienna University of Economics and Business Administration to examine the application of the Universal Service Directive and the Austrian telecommunications law. They produced a preliminary expertise [16] which supported our view to some degree, but was not conclusive.

In Germany, the new telecommunications law resulted in a voluntary commitment of the Deutsche Telekom to fund a three-year project for the implementation of a German relay center. We tried to come to a similar agreement with the Austrian telecommunications providers (as a lawsuit may last for years, with an uncertain outcome, a voluntary agreement would be much preferable); for this, we contacted the Austrian Regulatory Authority for Broadcasting and Telecommunications. They arranged several meetings with representatives of the different providers. Although we received some important information on broadcasting and connections, funding will probably only be provided if they are forced to do so by the law. This problem of assigning responsibility – as there are many competitors, no longer a state monopoly – has also been remarked on by the Nordic Forum for Telecommunication and Disability:

The liberalisation of the field of terminals and services has resulted in a difficulty of assigning any responsibility. The number of suppliers has been multiplied and the market is therefore likewise opaque. It has become easier for the suppliers to "hide in the crowd" with the result that the responsibility which formerly could be placed unambiguously through the political control with the monopolies today is lifted in reference to the free competition. The social responsibility, which the former companies with monopoly had because of their status of public service institutions, has more or less disappeared.

The consequence is that today it is far more difficult to assign the responsibility for the accessibility of new terminals and services for persons with disabilities. In a market of competition it is obviously difficult to get anybody to assume the responsibility of a non-profit field as aids for disabled [persons] or those services, which users with special needs require [17].

It may be helpful, however, that the Austrian government has decided to recognize Austrian sign language as a minority language on July 6, 2005. However, this amendment to the Austrian constitution states that the details are left to the individual laws. Therefore, any services which are not explicitly stated will have to be negotiated with the government. If the result remains unsatisfactory, a special organization, the Klagsverband, can take action. Mediation is obligatory; if this fails, the case can be taken to court.

4.5. National and international partners

Talks were initiated with companies which sell and/or develop multimedia software either for relay centers or for video conferencing. Some of these products were aimed at deaf and hard-of-hearing users, while others were mainstream solutions (mainly for intra-company communication); the former were preferred, because although the companies were much smaller, they were aware of the special needs of hearing-impaired people and had taken them into consideration. For example, a high-quality video is necessary for a sign language conversation[8], and the whole upper body of the communication partner must be visible, not just the face. We also discussed with existing relay services whether they were interested in building up a relay center in Austria, and if so, on which terms. Private telecommunications providers and the former state monopoly, the Telekom Austria, were asked about a possible cooperation.

As for the contents, we used our contacts throughout Austria to find possible partners: from the sign language interpreters' training at the University of Graz to special hospital

[8]According to the Nordic Forum for Telecommunication and Disability, video communication suitable for sign language needs at least 20–25 frames/second and a minimum resolution of 352×288 pixels (corresponding to the common intermediate format, CIF). The delay must not be more than 0,8 seconds, ideally less than 0,2 seconds, to make lip-reading possible [18].

departments for deaf people, we presented the concept of a relay center to institutions dedicated to the education, rehabilitation and/or well-being of deaf, hearing- and speech-impaired people. Other departments of our own university were interested in participating as well, e.g., the informatics department.

An international cooperation would be possible especially with other German-speaking countries, i.e., Switzerland and Germany (for instance, while the sign languages differ, text relay could be shared during the slower night hours). First talks have already taken place. Further cooperation with the neighboring countries of Northern Italy and Slovenia (the Alpen-Adria regions) are conceivable as well.

4.6. The concept of the Relay Center Austria

Based on what we learned and what experienced relay centers shared with us, we developed an integrative concept for the Relay Center Austria. What started out as a simple relay service, soon turned into a comprehensive service center for the target groups. This decision was motivated by two main factors. First of all, it was almost impossible to predict how many interpreters would be needed in the beginning – depending on the target groups' acceptance of the relay service, it was possible either to end up with interpreters twiddling their thumbs or being seriously overworked. In order to avoid this, we adapted a Swiss idea, where the communications assistants write the teletext for the deaf during the night hours, when there are fewer calls. Second, during our presentations we encountered many projects (e.g., SignTime, a Vienna-based project for sign language news for the deaf), where work had to be stopped because of a lack of funding. The Relay Center Austria, with its resources of multimedia technology and interpreters, could serve as a kind of umbrella organization and cooperation partner for such efforts. It might also be easier for the authorities to deal with a single service center.

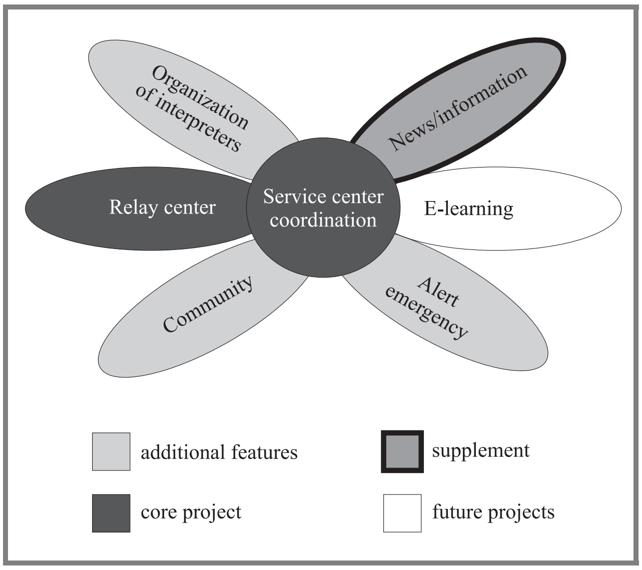


Fig. 4. Concept of a service center.

The relay center itself should be embedded in an Internet portal which offers other services as well. We developed six basic services of varying urgency; some of them are aimed at all the target groups, some only at the deaf (Fig. 4).

In accordance with a first feedback by the deaf associations, the relay service will be implemented first, followed by the other services. The order in which they will be implemented as well as the details will be developed in close cooperation with the target groups.

4.6.1. Relay service

This core service will consist of text and video relay in one, as we will use a product combining video, text and speech. Text will mainly be used by people who do not know sign language, but it may also be used to spell out a name or a difficult word or as a possibility to fall back on when the connection deteriorates. The user can freely choose which of these communication channels s/he wants to activate, based on their equipment or their personal preferences.

4.6.2. News and information in Austrian sign language

Barrier-free access to information is one of the most important issues today. Therefore the website of the relay center should offer news and information for deaf people, ideally in the form of sign language videos (with subtitles or additional text for non-sign language users). There could be a cooperation with institutions which already have this service, e.g., the Austrian Deaf Association[9] and a website with online information for people with special needs, BIZEPS[10]. There have also been talks with the Swiss "sign TV" Focus-5, which is run by deaf people[11].

An additional service could concern too difficult texts: Deaf people could send them to the relay center and get them back either in an easy-to-read version or signed by an interpreter.

4.6.3. Alert/emergency

As most alerts are in acoustic form only (e.g., fire alerts, civil-defense alerts, traffic news), some alternative has to be found for deaf people. One of the possibilities is an SMS alert, although this is not ideal because of the inherent drawbacks of SMS messages (no absolute reliability, character limit, speed, no direct contact, etc.). Again, the details have to be discussed with the target group.

Another problem concerns deaf or hearing-impaired people who need to contact an emergency number. There is, for example, a respective project in Sweden. The (partly negative) experiences are summed up in [19]. In some countries (e.g., Australia or the USA), text relay can be used to re-

[9] Cf. http://www.oeglb.at/
[10] Cf. http://bizeps.or.at/oegs.php
[11] Cf. http://www.focus-5.tv/

port an emergency[12]. Although video relay seems to be the perfect solution for sign language users, extensive tests are necessary to determine its applicability.

4.6.4. E-learning

The website should include an e-learning platform with courses aimed at both deaf/hard-of-hearing and hearing people. For the hearing-impaired, there will be courses with sign language as language of instruction, e.g., German, English and various computer courses (for instance, the European computer driving license/ECDL). The hearing should get the possibility to learn some sign language online, via courses and the database that is being built up by the Center for Sign Language and Deaf Communication[13].

4.6.5. Organization of interpreters

In Austria, few sign language interpreters work full-time. Therefore, it is often difficult to get one, especially if you do not book them some weeks in advance. In most of the Austrian provinces, there is no central organization, either; you have to call each interpreter individually, until you find one who is able and willing to take on the job. The relay center could help by keeping track of when each interpreter is busy. When an interpreter is needed, it would suffice to call the relay center and provide them with the details of the interpreting situation; the relay center would then do the rest and try to find a free interpreter (similar to Switzerland, where you simply fill in a form and send it to the Swiss relay center PROCOM[14]).

4.6.6. Community features

The website will also offer community features which should be managed by the target groups themselves. The details are to be discussed yet, but these features could include, e.g., a message board, chats, an electronic calendar of events, and user profiles (so that a search for other users with special hobbies or qualifications is possible).

5. Next steps

Due to modern communication and information technology, the relay center can consist of several locations in different provinces of Austria. This separation is also important for back-up purposes – if one location breaks down, the service can be switched over to another location. The administration will be centralized, however. The first location is to be in Carinthia, due to the involvement of the Center for Sign Language and Deaf Communication and the University of Klagenfurt.

As for the next steps, first of all, the necessary funding has to be secured from the government. There has to be an extensive cooperation with the target groups and the sign language interpreters, and the details of the proposed services have to be agreed upon. The infrastructure will be installed, followed by a trial period with the chosen software and hardware. In the meantime, the public needs to be informed about the Relay Center Austria – the hearing must be forewarned about possible calls from the relay center, else they might think it to be some elaborate hoax.

During the first phase, relay services should only be available for two provinces, to test the technology as well as the procedures and to deal with any problems. As soon as this limited service functions satisfactorily, it will be expanded to the other provinces and finally to the whole of Austria.

6. Conclusion

For hearing people, spoken language is the preferred means of communication. Because they are the majority, this is not going to change in the near future. Therefore, those people who are excluded from spoken communication must be provided with an alternative: in the case of phone communication, with a relay center.

Such a relay service does not only guarantee more independence to the deaf, hearing- and speech-impaired and make it easier for them to contact hearing people, but it has the same effect in the opposite direction as well: hearing people can call deaf, hearing- or speech-impaired friends, relatives, etc., directly. No longer is there the problem of faxes, e-mails or SMS messages which remain unanswered (because they did not arrive, were discovered too late or have not been understood).

There are additional bonuses, as the president of the Swiss PROCOM remarks: not only are there fewer prejudices against deaf people, because there are no mistakes in a written text, but a call by a relay center also serves to make hearing people aware of the problems the deaf face regularly in their everyday lives [20].

The integrative concept of the Relay Center Austria – not only a relay center, but a service and competence center for deaf, hearing- and speech-impaired people – makes sense both from an organizational and a business perspective. Furthermore, it allows for a better inclusion of the target groups into the hearing society. Not only does it improve their access to information, by removing or at least lessening the existing deficits like acoustic alarms or a lack of written language competence, but it also allows for e-learning – both with courses which are custom-made for the target groups and by expanding the number of sign language competent hearing people.

In order to achieve this, the following points need to be emphasized:

- No separation between vocational and private inclusion.
- A holistic view of the project (no isolated solutions).
- Integration of existing measures and initiatives.

[12] In the USA, the ADA requires all 9-1-1 centers to have a textphone, while in Australia, there is a special "106 Text Emergency Relay Service", cf. http://emergencycalls.aca.gov.au/ace.htm or http://www.aceinfo.net.au/Resources/Downloads/factsheets/pdf/015ers.pdf

[13] An online version used in the project "Sign-IT" can be found at: www.sign-it.at/ ("ÖGS-Lexikon").

[14] For the form, cf., http://www.procom-deaf.ch/procom/s/dolmetschdienst.asp

- Use of modern communication devices (without excluding older devices, true to the philosophy of universal access).
- Setting up of a competence center for the target groups.
- Support for any initiative coming from the target groups.
- Support for (higher and secondary) education of the target groups.
- Barrier-free access to important information.

The final aim is a business which is familiar with the special needs of the target groups.

References

[1] Nordic Forum for Telecommunication and Disability, "Nordic guide to text telephony", NFTH 2/2002, 2002, http://www.nsh.se/NFTH-Text_Telephony.pdf

[2] C. R. Steinhäußer, "Technische Hilfsmittel zur besseren Integration von Gehörlosen", Diplomarbeit Technische Universität Graz, 2000, http://www.taubenschlag.de/lernen/wissenschaft/steinhaeusser/technische_hilfsmittel.pdf

[3] "CSD interpreting online", http://www.csdinterpretingonline.com/index.html#

[4] F. Heinrichs, "Technische Vermittlungsdienste für Hörgeschädigte in Deutschland", Magisterarbeit Universität München, 2004, http://www.taubenschlag.de/lernen/wissenschaft/heinrichs/TechnVermDienst.pdf

[5] W. Andersen, "In which cases does videophone interpreting work good or not so well?", in *Report Nordic Conference for Video Phone Interpreters*, Castberggård, Denmark, 2003, Annex 2, pp. 10–13.

[6] R. Jacobsen and L. Green, "Videophone vs normal interpreting", in *Report Nordic Conference for Video Phone Interpreters*, Castberggård, Denmark, 2003, Annex 2, p. 18.

[7] R. Spencer, "Video relay interpreting trial final report", NRS DOC 030, Australian Communication Exchange, 2000, http://www.aceinfo.net.au/Resources/Downloads/VRI/2000/vri_trial.pdf

[8] Research Center for Sign Language and Communication of the Hearing-Impaired, University of Klagenfurt, Final document for the project "Preparing a European Deaf Network for Information and Communication", Klagenfurt, 2000, http://www.uni-klu.ac.at/zgh/eu_projekte/2000/

[9] Center for Døve, *Report Nordic Conference for Video Phone Interpreters*, Castberggård, Denmark, 2003.

[10] T. F. McCaul, "Video-based telecommunications technology and the deaf community. Research summary of the report", RC DOC 007, Commonwealth of Australia, Australian Communication Exchange, 1999, http://www.aceinfo.net.au/Resources/Downloads/VRI/1997/VRI_full.pdf

[11] P. Roe, "Comparative study of relay services 1991–2001", 2001, http://www.stakes.fi/cost219/Procroerelay3.doc

[12] M. Hilzensauer, F. Dotter, M. Maitz, M. Frühstück, A. Hopfgartner, J. Oberauer, and A.-M. Valentin, "Relay Center Austria: Bericht zur Projektvorbereitungsphase. Ein Projekt im Auftrag des Bundessozialamts", Klagenfurt, Zentrum für Gebärdensprache und Hörbehindertenkommunikation, 2005 (manuscript).

[13] "Equal measures: closing the accessibility gap", T. Shipley and J. Gill, Eds. 2005, http://www.tiresias.org/cost219ter/equal_measures/index.htm

[14] W. Jolley, "When the tide comes", in "Towards accessible telecommunications for people with disabilities in Australia: a discussion paper commissioned by the human rights and equal opportunities commission", 2003, http://www.hreoc.gov.au/disability_rights/communications/tide.htm

[15] Nordic Forum for Telecommunication and Handicap, "Regulation and financing of telecommunications services for disabled persons in the Nordic countries", NFTH 2/2004, 2004, http://www.nsh.se/NFTH-Regulation.pdf

[16] M. Holoubek and D. Damjanovic, "Gutachtliche Stellungnahme zur Frage der Finanzierung eines Telefonvermittlungsdienstes für Hör- und Sprachbehinderte im Rahmen des Universaldienstes gemäß §§ 26 ff TKG 2003", BGBl I 2003/70, 2004 (manuscript).

[17] Nordic Forum for Telecommunication and Disability, NFTH policy document. "Analysis of and strategy for NFTH's future work. Accessibility to the information society for disabled and elderly people", NFTH 1/2002, 2002, http://www.nsh.se/NFTH-Policy_document.pdf

[18] Nordic Forum for Telecommunication and Disability, "Nordic guide to video telephony and video relay service – for persons with impairments", NFTH 3/2002, 2002, http://www.nsh.se/NFTH-Video_Telephony.pdf

[19] J. Åslund, "Report summary. Test of SMS – chat for emergency number 112". Stockholm, Netlight Consulting, 2003, http://www.netlight.se/reports/112-SMS_Chat_Summary.pdf

[20] B. Kleeb, "Die PROCOM-Telefonvermittlung in der Schweiz – und ein Vergleich mit Deutschland". Presentation at the "Kulturtage der Gehörlosen" (Days of Deaf Culture), Munich, Germany, 2001, http://www.taubenschlag.de ("Wissenschaft" → "Referate").

Marlene Hilzensauer was born on November 9, 1967 in Klagenfurt, Austria. She studied Applied Linguistics and English at the University of Klagenfurt. After her Ph.D. graduation in 1994 she participated in several projects (English as a Second Language and Carinthian dialects). She joined the project "Austrian Sign Language" at the University of Klagenfurt in 1995 and has worked there ever since. The project developed into a subdepartment of the Department for Linguistics in 1996 ("Research Center for Sign Language and Communication of the Hearing-Impaired") and a Special University Department in 2004 ("Center for Sign Language and Deaf Communication"). Her main fields of work are preparing multimedia courses/CD-ROMs and teaching materials for deaf (e.g., for German, English) and hearing people (for Austrian Sign Language) as well as the realization of a relay center for deaf, hearing- and speech-impaired people.

e-mail: marlene.hilzensauer@uni-klu.ac.at
Center for Sign Language and Deaf Communication
University of Klagenfurt
Universitätsstraße 65-67
9020 Klagenfurt, Austria

Paper

Inclusion of sign language users via information and communication technology

Franz Dotter and Marlene Hilzensauer

Abstract— The paper addresses the situation of sign language users (mostly deaf people) in the context of inclusion as a political goal. For several reasons, there is often still some confusion with the terms of deaf and hearing impaired. In order to overcome this confusion, a survey is given over the needs of people who have a sign language as their preferred language as well as the needs of people who decide on preferring spoken language (mostly hard-of-hearing people). One should also doubt that the whole target group of people with disabilities in the hearing area consists of two separate groups only. Starting from the right of self-determination, the better solution seems to be the individual right of a person to choose any offers which are useful for her/him. As for other groups of people with special needs, ICT is seen as a big chance for improving their situation in terms of life and job chances. Several projects and experiences are reported.

Keywords— integration, inclusion, deaf, hearing-impaired, communication society.

1. Introduction

There are a lot of programmatic declarations and initiatives concerning "integration" or "inclusion" of people with special needs and there is also much "good will". In case of the "hearing impaired" taken as the comprehensive group of people with any hearing disability we have to deal with a lot of misunderstandings and even ideological positions which make the situation more different than with other groups of people with special needs. Therefore, in order to reach optimal results for the inclusion of deaf people, besides providing adequate technical solutions two main conditions have to be fulfilled. First, the situation and needs of the sign language users have to be clearly acknowledged (especially in comparison to other hearing-impaired people). Second, the intellectual, scientific, organizational, and political barriers for the necessary developments have to be brought to the attention of all decision-makers so that they can be removed.

The objectives of the paper are to clarify the specialty of the target group addressed – in contrast to the majority of the hearing-impaired – and to draw the necessary scientific, educational and organizational consequences from the given situation.

2. An unclear target group?

2.1. Acoustic perception and spoken language

Compared to salient acoustic phenomena like a strong pulse or a loud noise, the perception of a spoken language needs fine differentiation with respect to acoustic data. The differences in intensity between accented and non-accented syllables or words are big; the same is valid for intensity and frequency bands of different sounds. Therefore persons with a restricted hearing ability are in danger to miss some less salient sounds within words or some less salient parts of words, words themselves or even phrases. The perception is additionally influenced by environmental noise and the communication situation. There are a lot of hearing-impaired persons who can perceive spoken language in a relatively silent environment and concentrating on one communication partner, but fail to understand communication in a noisy environment or in groups of people speaking to each other.

If the perceivable linguistic information does not reach a certain threshold, the acquisition of spoken language may be negatively influenced. During their ontogenesis of language, the respective persons then acquire "gaps" which lead to a less than complete mastering of their mother tongue compared to subjects with normal hearing.

Many forms of a hearing restriction have no severe consequences and lead only to slight deficits concerning spoken language or communicative competence. The situation begins to change with increasing hearing loss and becomes completely different if the hearing capability of a person is not sufficient to acquire spoken language via the acoustic channel in the manner that hearing subjects do (this is the practical definition of deafness).

While many of the hard-of-hearing only demand adequate hearing aids in order to amplify acoustic phenomena, the acoustic channel is barred to the deaf so strongly that more or less all acoustic data have to be presented in a visual form in order to be accessible. This is also true for spoken language. As there are often problems with or restrictions on the use of written language (for the reasons, see below), practically all the deaf choose a sign language as their preferred language.

In the case of deaf-blindness, both channels, the acoustic and the visual one, are barred and language has to be learned via the tactile channel (sometimes in combination with residual abilities in the other two channels). Here a system of tactile perception of a sign language is recommended, accompanied by written language in tactile form.

2.2. Spoken and written language

2.2.1. The learning of language

Human beings learn language as a special behavior in the framework of everyday interaction. They acquire competence by "anchoring" the signs of language on the results of general cognitive processes. They learn languages in realizations which can be quickly and spontaneously produced and perceived. For hearing people the acoustic mode is the best one for this enterprise, for deaf people it is the visual mode.

Though it seems sometimes that written language represents language much better than the spoken one, we learn it the other way round, naturally. We have a good competence in spoken language already when we begin to learn the written variant of our mother tongue. Although we should not say that it is impossible for a young child to learn a language first by its written variant, we all know that this would be a very unnatural way (how would we connect writing activities to spontaneous interaction?).

2.2.2. Principal differences in spoken and written language

There is another important fact concerning language use and learning. Our everyday spoken language is not identical with the standardized written language. These two language forms have often different communicative settings and functions. Due to its production mode and our normal communicative behavior, spoken language shows some features we normally do not find in written language (we do not always "speak as we write"), e.g., different registers and styles, breaks and new beginnings, or additional "paralinguistic" information.

Written language is of great importance for all people in our culture (as the so-called information or knowledge society strongly relies on written information). Deaf people can only be included in the hearing society if they have all written information at their disposal.

There is also the need to offer written English as a third language – in a form specially adapted for deaf people – in order to make deaf people able to follow the international communication.

2.2.3. Inadequate education

Due to shortcomings in deaf education yet existent in several countries, deaf people often lack a sufficient competence in writing and reading. For hearing subjects, spoken language is the main basis for learning the written language; this pedagogical principle cannot apply to deaf people, however. As a consequence, those deaf brought up following this inadequate principle are not only cut off from acoustically offered information but often also from written information because they had no real chance to acquire a sufficient competence. This is the reason why many sign language users need help with written language in spite of the fact that it is already in the visual mode. A consequence of this situation is that some deaf do not dare to use "normal" means of traditional or electronic written communication like letters, fax or e-mail with persons they do not know.

2.2.4. The social situation of deaf people today

Large parts of the groups of deaf people still show typical properties of underprivileged societal groups. Their job chances are worse, their jobs are less qualified than in the average of the whole society; a vocational career onto higher levels of organization and management is more or less closed (they often use the metaphor of the "glass ceiling" for this situation). In rich countries they are often cases for the social welfare system instead of having the chance of showing their abilities; in poor countries the situation is still worse. No wonder that the suicide rate of deaf people is higher than average. It has to be stressed therefore, that this partially enormous reduction of life chances is not due to the deaf's attitudes or behavior or to their special disability, but it is due to an inadequate furtherance and education and the discrimination which follows from these.

Although there have been massive improvements in some countries or sectors, there is still a grave lack of access to information and communication from the perspective of the deaf as a specific user group within the information and knowledge society. Video telephony, interpreters and fax are currently very important for work, but are not offered everywhere or wherever necessary. The same is valid for relay services.

2.2.5. Adequate education

The goal of offering written language competence to the deaf comparable to hearing subjects can only be reached if their respective instruction is considerably improved. As deaf people cannot hear spoken languages well enough to learn them this way, they need a special form of bilingual education, where a sign language provides the linguistic and cognitive base and the/a national written/spoken language is "anchored" to the first system (as a parallel but second language).

The sign language they chose as their preferred language has to be systematically used as the language of their instruction and – where they want to have it – also as one language of communication. Their competence in a written national language as well as in English have to be brought so far that they meet the normal educational standards for hearing people. More hearing persons should be motivated to learn a sign language as a foreign language in order to help inclusion from this side.

The ICT will play an important role in this task. There is the need for special research and optimization of this sort of language learning.

2.2.6. Still dangerous: a false analogy

The situation of deaf people with written language just described above explains why an appealing analogy falls down. Our experience shows that blind persons can easily read if they are offered a tactile version of written language (Braille). The reason for this is: they had no severe problem to learn the spoken language and can therefore acquire written language from this base like other hearing subjects. In this case the so-called compensation hypothesis applies: if an individual cannot perceive data from a certain sensory channel, these data have to be offered in another, accessible channel.

For deaf people the analogy would be: as they are unable to access the spoken language in the acoustic channel, we simply have to offer them spoken language via written language and all problems should be solved. Because this solution does not respect the natural learning sequence (spoken to written language) and because we cannot produce and perceive written language as simply and quickly as spoken language, the analogy does not work. Written language is by far not the best choice to replace spoken language for deaf people. Therefore we have to acknowledge sign languages as the languages chosen by deaf people for everyday use and instruction.

2.3. Two frequent misunderstanding concerning "sign language"

2.3.1. Sign languages and fingerspelling

Fingerspelling means that we use some different handshapes (they may be one- or two-handed) in order to represent letters from the alphabet of our written languages. Though fingerspelling plays some role within sign languages, especially when unknown names or new words coming from written language are introduced in sign language communication, it is not identical with sign language.

2.3.2. Sign languages versus signed oral languages

"Signed (oral) languages" are systems which (in a strict sense) provide a morpheme-to-morpheme transposition of an oral language into the visual channel. By this process all structures of the respective oral language are preserved. From this point of view, "signed" English, French, or German, etc., can be used in two important "bridging" functions. They can be used for hearing parents of deaf children to make the beginning of a visual communication system less difficult. And they can be used to show the structures of the specific oral language to deaf children in bilingual education. It has been found that a "signed oral language" is an operative means for initializing visual communication. But it shows no features of an adequate use of the visual channel (including factors of linguistic economy) so that it remains relatively slow and complicated. Therefore, if an individual should develop a visual language of his/her own, it should be a sign language as soon as possible.

2.4. The target group reconsidered

2.4.1. Self-determinacy

A principle to be stated here first is that of self-determinacy. Everyone has the right to define the form of his/her life. Parents of a child with hearing difficulties have to decide for their young child, the child itself has to get more and more control over his/her language choices as it matures. Self-responsible adults have the right to get language and communicative contexts as they want them.

2.4.2. Sign language users as a minority within the hearing impaired

To describe the minority state of deaf people within the large group of the hearing-impaired, let us look at the numbers. The standard estimation of the number of deaf people (= primary sign language users) in any society is 0.1–0.2 per cent. The percentage of hearing-impaired people as a whole varies from country to country and depending on the criteria applied, but with about 6–15 per cent it is sixty to hundred times the number of deaf people. From this fact it is understandable that the use of the generic term "hearing-impaired" in political or scientific discussion often leads to the consequence that only the needs of the overwhelming majority within this group, the hard-of-hearing, are considered. This fact is the reason that many scientific publications simply ignore the fact that there is an important division within this group concerning self-identification and language needs.

2.4.3. Sign language users as a linguistic/cultural minority

Due to the use of a special type of (visual) language, deaf people most often identify themselves as a linguistic/cultural minority. Therefore we have the – perhaps surprising – situation that a group of people which is societally identified as having special needs, declares to be a minority with special language rights. To be "deaf" in this sense is therefore not a matter of any medical diagnosis but of self-identification.

In analogy to the ethical rules of working with "exotic" oral languages and cultures, deaf as native speakers of their sign language must get their full "language rights", including recognition of the national sign languages similar to oral minority languages – if they demand that. The consequences are, e.g., deaf people are primary candidates to give courses in "their" language (that follows from the "native speakers first"-principle). Interpreters have to be trained like those for oral languages. The deaf communities must have the possibility to (co)determine plans and work on the scientific analysis of sign languages, deaf education, and the development of their facilities.

2.4.4. An open and changing sign language user community

In earlier times, this community could be easily described as consisting of the deaf people themselves, those relatives and friends of them who used a sign language in order to communicate with them, and the interpreters. Like some other minorities, some deaf groups even looked closed for people coming from outside.

The traditional understanding of "disabled", realized, e.g., in special schools for the deaf, made segregation and discrimination possible. The special schools not fulfilling the demands of standard education for a life of inclusion in society hindered many parents to let their child attend such a school.

There are several factors which since at least one or two decades lead to an immense change within the sign language user community:

- The deaf communities open to the hearing society.
- Research into sign languages and deaf culture have positive results, e.g., for the acknowledgement of sign languages and the improvement of deaf education.
- More hearing people find a sign language attractive enough to study it as a foreign language.
- There is more understanding in societies concerning the needs and rights of people with special needs.
- The stigmatizing of sign languages as primitive communicative systems (i.e., non-languages) and of deaf people as "dumb" decreases.
- The rigid delimitation between "deaf" and "hard-of-hearing" begins to become weaker; more hard-of-hearing people see a sign language as an additional medium of communication and no more as something to be avoided.

All the points listed above have the increase of the sign language community and its prestige as their consequence. But there is also a strong factor which directs the development to the opposite: with the invention of the cochlear implant (CI), a big change began to take place. First late-deafened adults, then deaf-born children were implanted in order to overcome deafness. The portion of implants in the rich countries or in rich parts of a society reaches about 80–90 of these people.

While the implants help almost all late-deafened adults, the implantation of deaf-born children is not so successful. Most important is that a cochlear implant does not mean that the respective person will identify always and only with the hearing society. This identification depends on the quality of understanding spoken language they reach with the implant. As a consequence, the number of deaf people in the traditional sense may diminish, but the need for sign languages will remain. Some hard of hearing people with a severe impairment may use a sign language, at least partially or for certain purposes. We estimate the number of hard of hearing who would eventually rely on a sign language with about 0.1–0.2 per cent of the population.

2.4.5. The global perspective

It should be added, however, that this development is only valid for the rich countries. Starting from a number of about 6 000 millions of human beings, there are at least 6 millions of deaf people worldwide for whom adequate measures of inclusion could bring about a higher quality of life. But the great digital divide existing between rich and less developed countries requires adapted (and many non-ICT-) measures.

3. Technical aids for the perception of spoken language

Before turning to the needs of sign language users, we give a short overview concerning the aids for people who decide to use spoken language as their preferred language (mostly: the hard-of-hearing). They need all equipment which helps to strengthen the acoustic data in order to be sufficiently perceived for, e.g., understanding language. As an alternative or additional help, the acoustic data can be transformed into visual ones (i.e., the spoken language can be written down). Within limits, the systems transforming spoken language into a written form make the former accessible also to deaf people.

3.1. Systems amplifying acoustic stimuli

These are, in general, all amplifying systems, hearing aids and induction loops. As there is much information and discussion already established, we do not go into details here.

3.2. Systems transforming acoustic stimuli

Here we find two basically different solutions.

3.2.1. Systems of manual or automatic transformation of spoken language into a written representation

While the manual services (i.e., subtitling, note-taking or captioning) are rather traditional, the automatic transformation from the spoken to the written mode is rather new, because it needs speech recognition. The European Union (EU)-project VOICE [1] is an example for such a system. The automatic recognition has currently still a major limitation. It forces the speaker to use a special sort of speech behavior (e.g., separating all words from each other by small articulatory pauses). All the systems – manual and automatic – have another limitation: they are useful only for persons who know a spoken/written language well enough to decode the meaning. As shown in Section 2, this is not the case for a considerable number of deaf people.

3.2.2. Systems for the transformation of the acoustic stimuli into electrical ones for the inner ear

These systems, the so-called cochlear implants are now recommended especially for deaf people whose inner ear nerves are in order. They are generally working very well for people who had some experience of spoken language before becoming deaf; the rate of success in cases of deafness from birth being lower and the subject of some controversies [2].

4. Technical aids for sign language users

The access of deaf people to communication and information has the following aspects.

4.1. Aids for telecommunication between sign language users

Here the transfer of visual information is necessary. In principle, modern ICT makes such a transfer possible. But still the quality needed for understanding sign language poses some restrictions on the types of telecommunication means used. While a double ISDN line provides a sufficient data rate, slow telephone connections do not. As a consequence, video telephony or videoconferencing via Internet are impossible for sign language users with a standard telephone connection. Mobile video telephony is sufficient within areas of UMTS service only. In all cases, the costs of a sufficient line capacity are a major problem.

4.2. Aids for the communication between sign language users and hearing subjects

In this case, the most economic solution would of course be an automatic translation in both directions. But there is no complete automatic translation system for any spoken/written language to any other (even for written variants); the situation being even worse for minority spoken languages as well as sign languages. Therefore human interpreters are necessary, the costs of which strongly limit their use. A good solution is represented by the so-called relay centers [3], especially those who offer written (= text) as well as signed (= video) access to their services.

4.3. Aids for making spoken information accessible to sign language users

Two solutions are possible. The optimal one is real-time or delayed interpretation from spoken into signed language (for problems, cf. Subsection 4.2). This solution offers the information in the language preferred by the addressees.
The second solution is the transfer of a spoken language from the acoustic into the visual mode, i.e., into written language (as described above in Section 3).

4.4. Aids for making written information accessible to sign language users

Human interpreters are also needed if we want to have the information translated into a sign language. Concerning automatic transfer from written language to a sign language, the EU-project eSign (a further development of ViSiCAST) [4] shows the state of the art. It converts frequent (written) phrases into fixed forms of a sign language produced by an avatar. This is no translation, but an important step towards more accessibility. This project aims to provide sign language on websites, using avatar technology and is currently working on local government websites in Germany, the Netherlands and the United Kingdom. They claim that, although the technology is still in development, virtual signing can be even better than using videos of real signers: changing small sections of the signing can be done quicker and easier (e.g., for updates), and the videos can be downloaded faster and do not take up much space on Internet servers.
Naturally, the improvement of the competences of sign language users in written languages could unburden the deaf from some necessity to use interpreters too often. To reach this goal, intensive educational measures have to be set, however.

5. Regulations and standards

There are some EU directives as well as international and national laws against discrimination, on human rights, on equal opportunities, or on inclusion of people with special needs.
The laws for equal opportunities, e.g., formulate the right of deaf persons to use sign language in many contexts. This includes the right to get interpreting financed. But – as they concentrate on equal rights at the workplace – they ignore the fact that the right of using one's preferred language is not very valuable if one cannot learn this language following the standards of language acquisition (this includes the systematic use of the language in education and the respective education of teachers).

Concerning standards (cf. the web accessibility initiative (WAI) [5] or the European Telecommunications Standards Institute ETSI [6]), the needs of sign language users are not always sufficiently taken into consideration. Though some standards give adequate recommendations, namely to use a written language, simple written language or a sign language in order to transfer acoustic information into a visual one, these are presented as if they were equally applicable and useful. Sometimes sign language is even mentioned as a "non-text equivalent of text". The focus is then on "clear and simple" (written) language, which should benefit deaf persons. While this is at least partially true, most deaf people would prefer a sign language. Neither the connections between deaf education and different competence in written language are made clear, nor the language rights of deaf people are mentioned. Additionally, there is no clear differentiation between sign language users and other people with a hearing impairment. Practically all recommendations focus on assistive technology, without telling us explicitly that some forms of this technology – especially those important for languages, including translations – are still not available. Adaptive technology is an important instrument for the inclusion of deaf people. But it has to be developed to a state where it works reliably (which is not the case for sign language synthesis and recognition yet) and it must not replace personal social interaction between deaf or deaf and hearing persons. Additionally, the potential users must be offered the competence to handle hard- and software.
From our perspective, this situation is mostly due to the following facts:

- The people with special needs under discussion are not always invited to participate in technical standards development.
- It is not understood that the inclusion of deaf people has to be provided not only by assistive technology but often needs direct human interactions.
- The technology orientation of scientific programs: sometimes it seems that many technicians are not interested in a real inclusion of groups of people with special needs if mainly already established technology is used. They are more interested in developing new technologies. In other words: there are conflicts between economic interests and the interests of people with special needs. Practical examples of necessary tasks, using established technology, are the extensive installation of relay centers or the installation of video connections between deaf organizations nationwide and worldwide.
- The costs per individual signer: as sign language users need their visual languages, they cause costs like any other acknowledged language minority or even more, because they often need interpreters in order to communicate with the hearing society. As they are few in numbers and spread all over the countries, the costs per capita are quite high. This fact often prevents them from getting the adequate service.
- Lacking information about the needs of the deaf: hearing people without direct contact to the deaf community are often convinced that deaf people can easily compensate by reading written texts.

6. The contribution of new media to inclusion

The new media (NM) or information and communication technology (ICT) give us for the first time the chance to realize education material and measures adequately, sustainably and economically. Sign languages as visual languages can be presented in digital films, software tools allow different designs and a flexible exchange of materials, as well as the cooperation of various authors. We can save, transfer or exchange these materials if they are in a format which allows that. Templates, i.e., parts of software which allow the use of a concept (e.g., of a course or an exercise) within different contexts, will play a major role here.
Because the target group is so small, it is necessary not to realize one and the same material, e.g., in every EU country at high costs, but to use templates or distribute parts of the work and then exchange them. Such a practice means that only the detailed decisions concerning the material and the video production have to be done individually, but the software costs and some design costs only have to be financed once.
Furthermore, ICT allows expensive courses – often held only a few times – to be stored, modified and called up easily. A comprehensive offer of online courses also helps to realize the concept of a free and accessible basic education and training.
The ICT allows for more self-determined work of the students, e.g., by improving temporal and local accessibility, either in combination with a presence course or via online courses alone. The same is valid for joining formal or informal groups of learners. At least some of the material can be used for other target groups (immigrants, alphabetization, people with learning or speech disabilities) also or some can be taken from already developed products for these groups.
Useful solutions in the new media should show the following properties:

- a flexible structure, e.g., modules instead of fixed lessons which allow the users free navigation;
- a software with tool character which allows any author to design more modules or exercises without having programming competences;
- a cognitive orientation which motivates the learners to start from their own experiences, e.g., with language and strategies of behavior and information processing (e.g., analyzing, contrasting, systematizing, deducing).

7. Recommendations for ICT in the area of deaf and hard-of-hearing users

7.1. Self-determination and multiple choices

To recall the principle, every deaf or hearing-impaired person should be able to decide for themselves, whether s/he wants to identify with a signing or a speaking community (this includes free choice concerning their preferred language) or even wants to live in both communities.
Therefore ICT for hearing-impaired persons must allow many individual strategies of approaching information and language. Ideally, the users have to be offered a general ICT-menu within which they can decide what (combination of) media they use for their access to information and communication.

7.2. Recommendations

We can classify the hearing-impaired persons roughly in two groups (without postulating that every respective person will fall in one of the two following groups).

7.2.1. Spoken language users

This is the group of hearing-impaired people (mostly with a lighter impairment) who decide on spoken language as their preferred language. They can be helped by optimizing the auditive perception via hearing aids and by improving the visual channel as an additional source of information (e.g., lipreading for the perception of spoken language, presentation of important information both in the acoustic and visual channel) in the sense of multimodal access.

7.2.2. Sign language users

The other group are those people who decide on a sign language as their preferred language and therefore need a special bilingual way of education and training. Members of this group get their access to information and language(s) by using a sign language or other visual communication systems like "signed English".
In order to support the improvement of education especially of the second group (with some impact also on the first one), a lot of work is urgently required to provide sign language items and sign communication as well as to provide the interconnection of sign language with written and spoken language.
We have to develop more programs for computers which allow:

- parents to learn visual communication systems including sign language;
- parents to look for the adequate combinations of visual and acoustic means of communication for/with their child;
- parents and children to take adequate exercises for the development of visual and oral communication systems.

Naturally, ICT cannot substitute for everyday communication but it can provide inestimable tools for strengthening and broadening language competence and information processing. The point has to be stressed that communication skills are learnt by communicating with human beings, however.
To mention only a few features which an adequate ICT should display (for a detailed discussion cf. [7]):

- The computer system and its Internet connection must be able to run digitized videos in a sufficient size and good quality (e.g., 16–20 frames/s). The possibility of videoconferencing should be available.
- Learning programs have to include acoustic and visual versions, e.g., sign language, signed oral language, as well as written and spoken language for every item. That means: all information which is now available in the acoustic mode has to be transposed into the visual mode step by step. Programs and materials have to be offered for the early years, for preschool and school education, as well as adult education.
- Where automatic facilities do not work, there must be personal services present, like interpreter or relay services as well as educational interaction, counseling and tutoring in the classroom or via Internet.

8. Policy recommendations

The following policy recommendations should be taken into account:

- Sign language users must become a group specifically addressed by the authorities in all questions of accessibility, education and vocational training. Without always questioning the necessity or the costs, this group has to be provided with equal opportunities in education from the early years on and at the workplace.
- The authorities should install comprehensive interdisciplinary applied research and the use of modern ICT, e.g., building a "deaf research area" and a coordination network concerning deaf education (in connection with the representatives of the deaf people) in order to avoid doubling or repeating the same projects.
- Some money from scientific or social funds has to be given directly to the deaf in order to enable them to release calls for projects on their own (the same is valid for all other groups of people with special needs).

- The deaf users should take part in the development of new technologies, both in design and development.
- Access to information and communication services should not be regarded only as a hardware problem but as a multi-factor problem influenced by the following issues: the language used, costs, technical standards, contents and whether human services are necessary.
- As for the costs for the use of information, access to the basic services of the information society should be free, not only for the deaf.
- The communication needs of the deaf cannot be satisfied with functional equipment alone. They also need services and multiple options so that the deaf can choose the most fitting one, no limitation to a single solution.
- As for telecommunications, we suggest the development of a respective "deaf workplace". Such workplaces should exist at home, on the job and in some public institutions.

9. Examples from projects

Some concrete examples from our work can serve to illustrate what can be done in the area of deaf education.

9.1. The e-learning for the deaf

9.1.1. Internet as a potential source of new employment possibilities for the deaf

In this project of the Leonardo da Vinci programme, several courses for the deaf were developed and held as presence courses in the partner countries (Czech Republic, Austria, Flemish part of Belgium and United Kingdom), concerning computer skills (basic skills, Internet, web page design), the respective written national languages, (computer) typography and Internet-related job-seeking skills. The materials for this course include textbooks, partially accompanied by CD-ROMs.

An added bonus was the collection of standardized signs for grammar and computer terminology. The signs were discussed during meetings of deaf people from all over the country; the signs they agreed on were then filmed and stored in the form of video clips on CD-ROM. These signs provide a basis for the next courses so that there is no confusion because signs from different regions are used interchangeably. Interpreters can now draw on a unified terminology.

9.1.2. Courses in written German

In Austria, courses in written German for deaf people are especially important. Within a project aimed at introducing a new, officially recognized training for (deaf) teachers of sign language, the Center for Sign Language and Deaf Communication (ZGH) developed a (written) German course for future teachers. The online modules of this course will allow the students to look up information about German grammar and to practice it by doing exercises on their own. We believe that it is best to use the target language, i.e., written German, as a basis, supplemented with sign language videos as a ubiquitous help, if the users have problems understanding the text. The deaf can first try to read the text on their own, without additional help. If any of the words are too difficult or if the meaning is not clear, a help function explains everything in sign language.

The students can do the exercises in whatever order they like; there is no fixed course structure. The user simply chooses a subject which is of interest to them, i.e., the present tense, and can then read the respective grammar explanations (or watch the accompanying videos in Austrian sign language = ÖGS) and do the various exercises. Those are taken from an exercise pool, using a random generator, so that the students will get new exercises each time they access the course. There are three levels of complexity: beginners, intermediate and advanced. Normally, the program will show all three levels one after the other for each exercise type, but the students may skip any of them if they do not want to do them or do not feel competent enough.

For the easier levels, feedback is sometimes inherent; for example, with drag-and-drop exercises, wrong choices will not stay in place. There is also a special button which allows the students to call up the correct solution; with another click on this button, they can repeat the exercise. In the cases where the feedback is explicit, this takes the form of red (wrong) and green (correct) boxes around the respective word, phrase, etc., or of a green hook and a red "X". In this way, mistakes are made visible unobtrusively; but there are no grades, so that the deaf students are not feeling as if they were back at school again (this was an express wish of the deaf participants).

9.1.3. Sign On

The EU-project *Sign On – English for Deaf Sign Language Users on the Internet* [8] is a cooperation of Austria, Finland, Iceland, the Netherlands, Norway, Spain, and the United Kingdom (October 2004 – September 2007). It is aimed at deaf people who use a sign language as their first or preferred language and are not complete beginners, i.e., they are e-mail and Internet literate and have a basic competence of English.

The goals of the project are to teach deaf people some skills, strategies and the confidence they need to exchange e-mails in written English and to read information on English language websites (one of the target groups are deaf researchers attending international conferences and wanting to stay in contact with people from other countries afterwards); sign bilingual learning materials (ten lessons in written English, national sign languages and international sign) will be produced, both for individual study and classroom situations.

9.2. Learning a sign language

For the deaf, it is important that as many people as possible learn sign language. This concerns not only deaf children of deaf or hearing people and their hearing relatives and friends, but also other hearing people. At the University of Klagenfurt, we have noticed a growing interest for learning Austrian sign language among the students. Some believe that this may help them with their future career (e.g., students of pedagogy), while others have deaf friends or relatives. However, there are lots of students, who are learning it out of pure interest only. We believe that immersion is the best method for learning a language, so Austrian sign language is the language of instruction; only during the first lesson is an interpreter present if there are any questions.

Naturally, sign language competence can best be acquired by interacting with native signers. If this is not possible or for repeating what you have learned at home, the second-best way is using sign language videos. Pictures (line-drawings or real photos), even with little arrows to mark the direction of the movement, cannot capture the reality of the sign; although they may serve as a memory-aid, if one has already learned the respective sign.

At the ZGH, e-learning materials for courses have been produced over the last few years, replacing the original books and VHS videos, which were not really comfortable for the students (imagine, for instance, searching for a certain sign by forwarding and rewinding the tape).

9.2.1. Austrian sign language course "ÖGS 1" on CD-ROM

Although the students in our sign language courses are very motivated, they had some difficulty remembering the signs and the grammar over a longer period of time and they kept asking us for some study tool. To fulfill this request, we developed a course on CD-ROM together with a local software firm.

The first six lessons of the original course used at the university were adapted for a multimedia course (the next six lessons are to follow in a new project). The exercises remained mostly the same, while the PC took on the part of the teacher. True to our philosophy, only native signers appear in the videos. Mostly, sign language is used (e.g., for the feedback); only the instructions and grammar explanations are in written German, but this will also change with the second part of the CD-ROM.

The six lessons all have the same structure: dialogues illustrating the main topic of the lesson, e.g., describing people, are followed by exercises and games. The main grammar points are explained together with videos of sample sentences. Notes on deaf culture are intended to prevent hearing people from making common faux-pas, e.g., politely walking around two deaf people instead of crossing between them. Individual signs can be looked up at the end of each lesson or in an overall alphabetical list. Details of the contents are presented in [9].

The exercises and games often take the form of drag-and-drop or matching videos and pictures/translations. Most of the exercises use a random generator: the pictures, etc., are either completely different or at least will appear in a different order, so that the students cannot simply memorize the correct solution. Feedback is sometimes automatic (with drag-and-drop, wrong choices revert to their original place) and sometimes signed (the videos are chosen by a random generator to avoid boredom).

There is no rigid structure to the course: although complete beginners had best do the lessons in the correct order, they can also choose to pick out the parts that interest them most.

The course is not intended as a stand-alone product, although it can be used as such. We recommend to attend a presence course as well or to have at least some personal contact with deaf people.

The reaction to the CD-ROM was very positive, and it even won a prize, the *Europäisches Sprachensiegel 2004* (European Language Quality Seal). As it was programmed without static parameters, its framework could be used for other sign languages (or even spoken languages) by exchanging the videos and some of the data.

9.2.2. Sign-IT

This was a two-year project funded by the Austrian Federal Ministry for Education, Science and Culture; it was carried out by the ZGH in cooperation with the University of Graz (sign language interpreter training) and the higher education institution "Joanneum", also located in Graz [10].

Because sign language is taught in Austria only at two universities (Graz and Klagenfurt), not all the people who would be interested in learning it can attend the presence courses. The aim of the project was to develop e-learning for students of Austrian sign language so that they could also learn or repeat it individually. Moreover, by making ÖGS available to any interested person, the project would also serve to raise awareness of sign language as a minority language.

The e-learning part is still a pilot and consists of an online e-learning platform, containing courses on deaf history and deaf culture and allowing communication between teachers and students, a CD-ROM for individual study (mainly aimed at interpreters; containing structured lessons with exercises and vocabulary) and an online database.

9.3. Tools for developing courses

9.3.1. The lesson manager

Originally, this tool was used for the production of the CD-ROM "ÖGS 1", but it can be used independently. It allows the linking of sign language videos, glosses and written text.

Imagine, for instance, that you want to link a signed dialog with its translation. After loading a video file into

the program, it can be split manually into smaller parts, e.g., phrases or sentences. Normal control elements are used. The user can work either at normal speed or choose to view the file frame by frame. Each of the pieces can then be linked to the name of the signer, the respective gloss(es) and the translation.
A preview function allows the user to check that the linking is correct; if not, changes are possible (the latest version of the lesson manager allows not only the exchange of glosses and text, but also the reassigning of beginning and end points of the video).
The program is in German and English; but it can easily be adapted for other languages with a built-in function which requires only the translation of the German/English terms into the new language.

9.3.2. The course editor

For the multimedia course in German, we teamed up with the Austrian Research Center Seibersdorf to develop an electronic tool called "Course Editor". With this tool, exercises for courses can be created and the finished course then administrated. The design of the so-called masks (templates for the exercises) can be individually changed, according to the designer's wishes. From the pool of finished exercises, any one (also those designed by another person) may be copied and used as basis for new ones; those can then be added to the pool in turn. Possible exercise types include common ones like multiple choice, cloze tests, complementing or merging sentences, matching text and pictures. All in all, there are currently 15 masks.
For sign language courses and language courses that want to show the contrast between spoken/written languages and sign language, new exercises using videos (e.g., matching signed sentences with their translations) will have to be defined in a new project.
Although the course editor was originally intended for a German course, it is not limited to a single language. As the course editor can be filled with whatever exercises and videos the designer chooses, it can easily be adapted for different languages, by simply swapping the contents. The exercises can also be tailored by the instructor according to the needs of the respective group of students.

9.3.3. The Klagenfurt database for sign language lexicons

The first version of the database was developed for a project aimed at implementing a special server for deaf people in 1997/98, funded by the Austrian Federal Ministry of Science and Transport. This original database has since been revised and extended several times; currently, we are working on a completely web-based version.
The database can be used for signs from different sign languages. For Austria, signs from all the provinces are entered[1]. Not only formational properties like phonetic-phonological and sublexical ones can be described, but also morphosyntactic and semantic ones. Users can search for single entries, properties of signs or via semantic fields. The main characteristics are the iconic arrangement of sign components and the open sets of categories/parameters and category values (new categories and values can be added at any time, if necessary). Details are presented in [11].

[1]The database can be found in [10] under the heading "ÖGS-Lexikon".

9.4. Barrier-free access to information for deaf people

For discussions of the situation of deaf and hearing-impaired people with respect to the Internet cf. [12, 13].

Barrier-free access to the Internet. One of the requirements of our modern communication society is access to information on the Internet. A lot has already been done for barrier-free access, but sign language is usually seen as only a desirable addition, not a top priority (see Section 5).

However, there are some initiatives by European governments to amend matters: in 2004, the German Federal Ministry of Health and Social Security supplied some of its web pages with additional sign language videos [14]. Not only did they receive enthusiastic feedback, but an accompanying survey indicated that the users would like more such videos. From all the people interviewed, only 10% understood the contents in written form only, more than 89% needed sign language in order to really understand the texts [15].
A similar venture has been started in Austria in spring 2005. There is a special website managed by the Austrian authorities which deals with e-government [16]. Citizens can get information from the various ministries, print out forms, etc. The project led to some information texts in ÖGS, but most of the written texts still have to be translated after the end of the project.
There are some websites – mostly maintained by deaf associations or institutions/organizations which work with the deaf – which show best practice. An excellent example of a bilingual website is "Sign Community" of the British Deaf Association [17], where users are free to choose between British sign language and a text version.
As for news and information, a European example of good practice is the Swiss "Focus-5 TV" [18], which is run by deaf people and covers many subjects, ranging from the deaflympics to films for children. This excellent website can really be compared to "TV for deaf people". In Austria, the "Österreichischer Gehörlosenbund" (Austrian Deaf Association) offers news in the form of video clips in Austrian sign language and text on their website [19]; since January 2005, they also cooperate with BIZEPS (an association which offers advice to people with special needs and their relatives), which publishes chosen bilingual news on their website "BIZEPS-INFO" [20] .
Unfortunately, such websites are few and far between. Although video technology has come a long way, producing sign language videos is still a time-consuming and therefore expensive process. Until there are new technological

developments, sign language will probably remain limited to those websites which take a special interest in the deaf community and maybe a few official websites.

10. Final thoughts

There are big changes going on in so different areas as economy, technology, education, inclusion. It is very probable that there are many relations between these changes; but not all developments point to the same direction: the development of technology offers many new perspectives on inclusion and participation on the one hand. On the other hand a so-called economical compulsion serves as an argument to restrict or give up, e.g., inclusive developments in social affairs, because some ideologies tell us that we do not have the money for them. For the sake of persons with special needs we have to intensively point out their right for a full participation in society, though. This participation cannot undergo evaluation from a simplistic thinking of effectiveness in profit terms. From a more elaborate economic thinking which includes social costs and an ethics of welfare for all, inclusion is cost-effective. We researchers in this area have to stress the obligation of our societies to offer equal life chances. Progress of sciences and technology will allow the realization of that goal in any case.
Concerning the field of hearing impaired persons with its strong internal differentiation technology in combination with the ethics of inclusion could allow for the offer of all useful solutions of inclusion and participation as well as for the free choice of the individual person from these offers.

References

[1] VOICE project, http://voice.jrc.it/home_en.htm

[2] F. Dotter and I. Okorn, “Austria’s hidden conflict: hearing culture versus deaf culture”, in *Many Ways to be Deaf*, L. Monaghan, C. Schmaling, K. Nakamura, and G. H. Turner, Eds. Washington: Gallaudet University Press, 2002.

[3] M. Hilzensauer, “A way of integrating deaf, hearing- and speech-impaired people into modern communication society”, *J. Telecomm. Inform. Technol.*, no. 2, pp. 32–40, 2006.

[4] Essential sign language information on government networks (eSign), http://www.sign-lang.uni-hamburg.de/eSIGN/

[5] Web content accessibility guidelines (WCAG) 2.0 (draft) of the web accessibility initiative (WAI), http://www.w3.org/WAI/intro/wcag.php

[6] European Telecommunications Standards Institute (ETSI), http://www.etsi.org

[7] Research Center for Sign Language and Communication of the Hearing-Impaired, University of Klagenfurt, Final document for the project “Preparing a European Deaf Network for Information and Communication”, Klagenfurt, 2000, http://www.uni-klu.ac.at/fzgs/eu_projekte/2000/engl

[8] Sign On, http://www.signon.fi

[9] M. Hilzensauer, “CD-ROM on Austrian Sign Language: Course OEGS 1, Lesson 1-6”, in *Computers Helping People with Special Needs. Proceedings from the 9th International Conference, ICCHP 2004. Paris, France, July 2004*, K. Miesenberger, J. Klaus, W. Zagler, and D. Burger, Eds. Berlin, Heidelberg: Springer, 2004, pp. 1099–1106.

[10] Sign-IT, http://www.sign-it.at

[11] K. Krammer, E. Bergmeister, F. Dotter, M. Hilzensauer, I. Okorn, R. Orter, and A. Skant, “The Klagenfurt database for sign language lexicons”, *Sign Lang. Ling.*, vol. 4, no. 1-2, pp. 191–201, 2001.

[12] M. Albiez, “Möglichkeiten und Grenzen der Internetnutzung für hörgeschädigte Menschen. Unter Einbezug einer Praxis-studie”, Schriftliche Hausarbeit zur Ersten Staatsprüfung für das Lehramt an Sonderschulen, Universität München, 2000, http://www.taubenschlag.de/lernen/wissenschaft/albiez/albiez.pdf

[13] S. Klotz, “Vergessene Zielgruppe des Internets? Gehörlose und Internet – Analyse der Bedürfnisse einer Zielgruppe und Schlulfolgerungen am Beispiel Berlins”, Potsdam, 1998, http://www.taubenschlag.de/klotz/arbeit.pdf

[14] Bundesministerium für Gesundheit und Soziale Sicherung (BMGS), http://www.bmgs.bund.de/deu/gra/gehoer/ghv_start.cfm

[15] “Gebärdensprach-Filme im Internet”. Eine Umfrage vom BMGS und Gebärdenwerk im Zeitraum von 15. Januar – 20. Februar 2004, http://www.taubenschlag.de/kommunikation/gebaerdenwerk/umfrageeregnisse.pdf

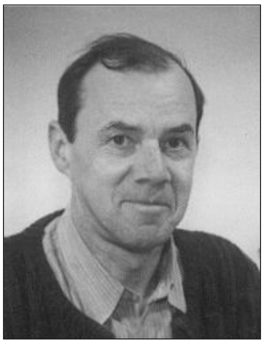
[16] HELP-Service for citizens, http://www.help.gv.at

[17] Sign Community, http://www.signcommunity.org.uk

[18] Focus-5.tv., http://www.focus-5.tv

[19] Österreichischer Gehörlosenbund (ÖGLB), http://www.oeglb.at/

[20] BIZEPS-INFO, http://www.bizeps.or.at/oegs.php



Franz Dotter was born in 1948 in Salzburg, Austria. He studied German language, literature, mathematics and philosophy. He received his Ph.D. in 1975. Since 1973 he has worked at the Institute for Linguistics and Computerlinguistics of Klagenfurt University. He habilitated (on iconicity in syntax) in 1990 and then worked as Associate Professor for general linguistics. Since 1996 he has been the Head of the Centre for Sign Language and Deaf Communication. His main interests focus on typology and cognitive linguistics, sign languages, sociolinguistics of politics and minorities, and text/discourse analysis.
e-mail: franz.dotter@uni-klu.ac.at
Center for Sign Language and Deaf Communication
University of Klagenfurt
Universitätsstraße 65-67
9020 Klagenfurt, Austria

Marlene Hilzensauer – for biography, see this issue, p. 40.

Paper

Implementation of information security management system in the small healthcare organization

Jiří Tupa and František Steiner

Abstract— The contribution describes the scope and main subject designed within DIGI-Q project. The paper contains results from subprojects of information security management system (ISMS) implementation, managed by students of DIGI-Q course. Very interesting simple risk assessment method and risk management and their application within in small healthcare organization were developed. Criteria and procedures accepted are described.

Keywords— information security management, personal data protection, risk analysis.

1. Introduction

The subproject solved problems with certification of data management system (DMS) in the private healthcare organization. The objective of this certification is to obtain licence certification mark GoodPriv@cy. The GoodPriv@cy is a certification service available for all IQNet clients interested in the protection of the company information data. Data protection and privacy (DPP) are becoming increasingly significant quality factors in business. It strongly affects the trustworthiness of a company or a public organization. The IQNet GoodPriv@cy specification integrates data protection and related information security requirements in a data protection management system. It supports an organization to manage its data protection and information security aspects proactively and efficiently.

The GoodPriv@cy label testiest that the authorised user: maintains a functioning data protection management system, meets statutory and contractual requirements for data protection and the related information security, continually improves the processes relevant to DPP.

The GoodPriv@cy label allows private and public organizations to document objectively and communicate effectively their own DPP performance vis-à-vis its customers and stakeholders. It's the way to safeguard reputation especially in healthcare public and private area.

The GoodPriv@cy certification can be used for small and medium healthcare organization in relation with quality management system. Example is shown by this paper. The attention is focused on a security of health information system contains medical documentation and sensitive personal data.

Implementation of DMS and its certification according to standards GoodPriv@cy can be way to information security management system (ISMS) building, especially in the health service. Why we are interesting about ISMS? Information is the lifeblood of all organizations and can exist in many forms. It can be printed or written on paper, stored electronically, transmitted by mail or by electronic means, shown in films, or spoken in conversation. In today's competitive business environment, such information is constantly under threat from many sources.

These can be internal, external, accidental, or malicious. With the increased use of new technology to store, transmit, and retrieve information, we have all opened ourselves up to increased numbers and types of threats. There is a need to establish a comprehensive information security policy within all organizations. You need to ensure the confidentiality, integrity, and availability of both vital corporate information and customer information.

What it is ISMS? An information security management system is a systematic approach to managing sensitive company information so that it remains secure. It encompasses people, processes and IT systems. For example ISO has published a code of practice for these systems – see ISO/IEC 17799 [6].

The personal data are the most important information in the healthcare organization. The goal of ISMS is the patient privacy protection especially privacy protection of sensitive personal data. Healthcare organization use information systems and information technologies for health documentation. This documentation contains sensitive data about all patients. Management and all staffs have to secure and develop system its protection. The GoodPriv@cy certification and audit is the way for good relationships between healthcare organization and patient.

The requirements for personal data protection are described by legislation documents: Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (directive on privacy and electronic communications), Directive 97/66/EC of the European Parliament and of the Council of 15 December 1997 concerning the processing of personal data and the protection of privacy in the telecommunications sector.

For example the Czech Republic and Slovak Republic passed the acts about personal data protection, which are harmonize with legislation of European Union law. These are act No. 101/2000 in the Czech Republic and act No. 428/2002 in the Slovak Republic. All type organizations have to respect these acts for processing of personal data.

2. Subproject description

The subproject was solved as a student work of DIGI-Q course. The DIGI-Q course was the part of project (Quality and On-line Confidence in SMEs e-Business Processes) is funded by the European Commission under contract number IST-2001-38157 of the Information Society Technologies (IST) programme 2002, key action II.1.3. The objective is to develop an ad hoc action of training and awareness for helping SMEs in improving and certifying their business and e-business processes. This will allow to increase the SMEs capabilities in being more competitive in the digital economy age.

Within this project, an advanced course is being developed to provide training to professionals working in SMEs, or in companies working for SMEs. In Czech Republic training course was performed in the Czech Republic by CQS – The Czech Association for the Quality System Certification DIGI-Q Consortium member.

The training course was two parts. First part – classes were given in "weekend blocks" and the second part – student works was solved six month as DIGI-Q subproject. Our subproject was entitled Data Management System Audit according to GoodPriv@cy and was focused on the healthcare area (Mr. Tupa and Mr. Steiner were trained by the course too and Mrs. Šebestová was tutor of training course and subproject). The subproject results were expected:

– audit of data management system;
– GoodPriv@cy certification.

The project results were obtained:

– design of audit checklist according to personal data protection acts;
– design of risks analysis methodology;
– implementation of risks analysis, suggestion how to decrease the risks;
– external audit of DMS according to national law.

The design of risk analysis methodology is the important part for ISMS. Risk analysis methodology was used for risk assessment and for risk management. The methodology was designed accordance with the requirements for personal data protection and according with the results of personal data protection system audit (GoodPriv@cy audit). This audit was made by external audit team.

The GoodPriv@cy audit was realized at a private healthcare organization by external audit team. The authors of this paper were member of audit team. Company was represented by director, quality manager, IT administrator. Audit was performed the way how the organization:

– a formulated and implemented data protection policy;
– an operational and documented DMS;
– compliance with all legal or contractual data protection requirements;
– provision of the information security by appropriate organizational, staff and technical measures;
– effective control and monitoring of processes;
– evolution and continuous improvement of data protection and privacy.

The review of fulfillment requirements with national law was benefit of certification. This act requires execution of external audit of DMS.

3. Implementation of ISMS for personal data protection

Requirements for fulfil of GoodPriv@cy mark is building the security management focused on the personal data protection. For that reason the standards for ISMS (for example standards ISO/IEC 17799 [6]) can be applied adequate for data management security. The ISMS system respected new approach for quality management and the standards for quality management ISO 9000 [7] and ISO/IEC 17799 [6] are harmonised.

The healthcare organization has quality management system (QMS) in accordance to standards ISO 9000:2005 [7] and the standards and documentation for personal data security was implemented to the QMS too.

General requirements for ISMS on the organization are: develop, implement, maintain and continually improve documented ISMS, in this case DMS is focused on the sensitive personal data protection within the context of the organization's overall business activities and risk. For the purpose

Table 1
PDCA description for ISMS

PDCA	Description
Plan	Establish security policy, objectives, targets, processes and procedures relevant to managing risk and improving information security to deliver results in accordance with an organization's overall policies and objectives.
Do	Implement and operate the security policy, controls, process and procedures.
Check	Assess and, where applicable, measure process performance against security policy, objectives and practical experience and report the results to management for review.
Act	Take corrective and preventive actions, based on results of the management review, to achieve continual improvement of the ISMS.

of this standard the process used is based on the PDCA model shown by Table 1 and Fig. 1.

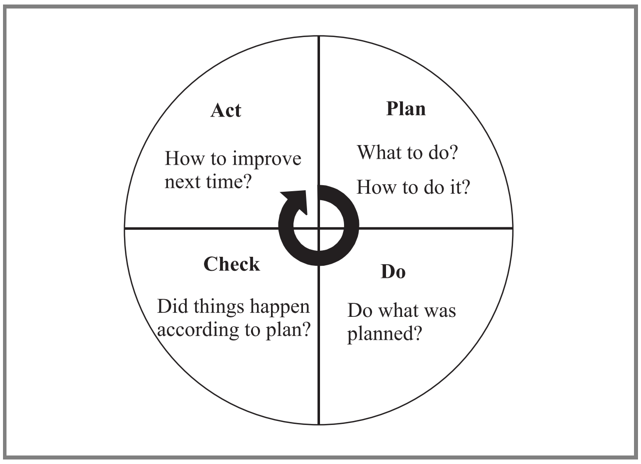


Fig. 1. The PDCA model applied to ISMS process.

The steps for establish the ISMS are:

- define the scope of the ISMS in terms of the characteristics of the business, the organization, its location, assets and technology;
- define an ISMS policy;
- define a systematic approach to risk assessment;
- identify the risks;
- risks assessment;
- identify and evaluate options for the treatment of risks;
- select control objectives and control for the treatment of risk;
- prepare a statement a applicability.

The next steps in the life cycle of ISMS at the organization are:

- implement and operate the ISMS;
- monitor and review the ISMS.

The important attributes ISMS are documentation (control of documents, control of records), establish the management responsibility (management commitment, resource management, training, awareness and competency), management review of the ISMS (review input and output and internal ISMS audit) and ISMS improvement (continual improvement, corrective action, preventive action) [1].

4. Risk analysis

Risk analysis is most important phase of ISMS establishment. That is process which compares assessed risks with benefit and/or price of possible security control. Standard ISO/IEC TR 13335 [8–10] defines four possible ways of risk analysis. We chose informal access. Advantage of this access is speed and financial modesty.

Base of risk analysis is fulfilment of following activities:

- threats identification;
- estimation of threat likelihood;
- identification of assets (process);
- rating of assets (process);
- determination of vulnerabilities;
- calculation of expected losses at threat impact;
- evaluation of risk analysis.

The interrelationships in risk management are shown in Fig. 2. This diagram helps us to tumble to risk analysis [2].

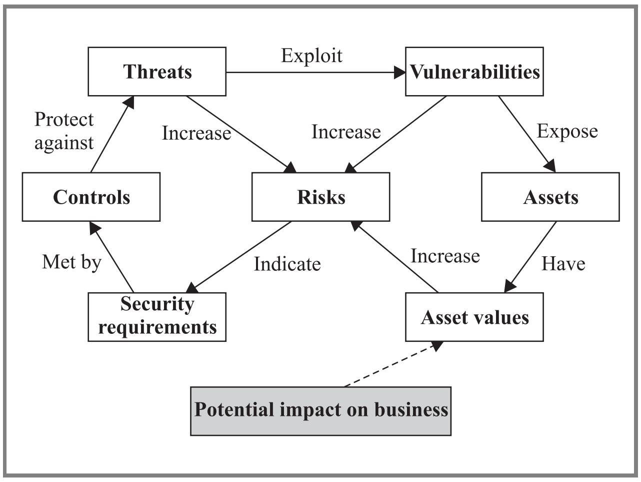


Fig. 2. Diagram of interrelationships in risk management.

The organization can be used standard methodology for example COBRA, CRAMM, etc., but this standards can be complicated and expensive, especially for smart and middle enterprises. In other case organizations can design own risk analysis methodology which will be easy and clear for the responsible management and fulfil the standards for ISMS.

The second case will be recommended for small and middle healthcare organization too.

4.1. Risk analysis key terms

The risk analysis key terms [3, 4] are given below.

Risk. The chance or likelihood of an undesirable event occurring and causing loss or harm. Note that the key element of risk is uncertainty, without which, there is no "risk".

Risk analysis. The process of gathering and analyzing risk-related information in the preparation of a risk assessment.

Risk assessment. A detailed articulation of the risks associated with the information assets and supporting ITC resources at risk, threats that could adversely impact those assets, and vulnerabilities that could allow those threats to occur with greater frequency or impact.

Threat. A potentially undesirable event that could result in loss or harm. The experience of a threat event and its measurable loss or harm is distinct from potential threat events and associated estimates of loss or harm. The aggregation of threat-event experience data provides the basis for estimating expected threat-event loss or harm in the future.

Vulnerability. A lack or inadequate application of a safeguard or control that allows a threat event to occur with greater frequency or impact.

Probability. Measures the chance or likelihood of an outcome or event occurring within a finite universe of possibilities or time, from zero (no chance) to 1.0 (certainty).

Uncertainty. The central issue of risk and risk metrics, reflected as the level of confidence, from zero to 100 per cent, that the associated numbers – and derived results – are credible and useful. Failure to integrate uncertainty into risk analysis/assessment approaches substantially reduces the credibility and utility of their results.

Quantitative risk analysis. It employs two fundamental elements; the probability of an event occurring and the likely loss should it occur. Quantitative risk analysis makes use of a single figure produced from these elements. This is called the "Annual Loss Expectancy (ALE)" or the "Estimated Annual Cost (EAC)". This is calculated for an event by simply multiplying the potential loss by the probability. It is thus theoretically possible to rank events in order of risk (ALE) and to make decisions based upon this.
The problems with this type of risk analysis are usually associated with the unreliability and inaccuracy of the data. Probability can rarely be precise and can, in some cases, promote complacency. In addition, controls and countermeasures often tackle a number of potential events and the events themselves are frequently interrelated. Notwithstanding the drawbacks, a number of organizations have successfully adopted quantitative risk analysis.

Qualitative risk analysis. This is by far the most widely used approach to risk analysis. Probability data is not required and only estimated potential loss is used. Most qualitative risk analysis methodologies make use of a number of interrelated elements:

- threats: these are things that can go wrong or that can "attack" the system; examples might include fire or fraud; threats are ever present for every system;
- vulnerabilities: these make a system more prone to attack by a threat or make an attack more likely to have some success or impact; for example, for fire vulnerability would be the presence of inflammable materials (e.g., paper);
- controls.

These are the countermeasures for vulnerabilities.

There are four types:

- deterrent controls reduce the likelihood of a deliberate attack;
- preventative controls protect vulnerabilities and make an attack unsuccessful or reduce its impact;
- corrective controls reduce the effect of an attack;
- detective controls discover attacks and trigger preventative or corrective controls.

4.2. Risk analysis methodology description – case study

Results of GoodPriv@cy audit made the base for design of the risk analysis methodology [5]. The methodology combines the basic principles of quantity risk analysis with qualitative risk analysis. The principles of failure modes and effects analysis (FMEA) analysis were used for the methodology design too.
The basic ideas of this methodology are:

1. The risks are generated by the processes in the organization that it means that we can search the risks in the processes which processing with the private information about patients and about organization (transfer, record, liquidation, storage, etc.).
2. The processes are part of an object in the organizations. An object makes the assets of organization and contains the processes which processing the information.

The risks are defined and described by individual objects for processes. For each process can be identified one and more risks. The result is defined by number which means risk factor of object. Average of all risk factor determines risk factor of all objects in the healthcare company. The results of risk analysis are recorded and summarized in the risk analysis forms (Tables 2 and 3).
The risk analysis form (Table 2) uses those important terms:

- *Process* – activities in the scope of object.
- *Risk* – description of risk.
- *Risk code* – No. of risk for risk inventory.
- *Risk assessment* – risk assessment contains three date:
  - *consequence* (*C*),
  - *likelihood* (*L*),
  - *risk factor* (*RF*).

Next columns in the form (Table 2) record the current safety precaution and control method of process (internal documentation for process control).

Table 2
Risk analysis form – example of object risk analysis

Object (asset):	**Information system**						**(Object code 0.1)**
Process	Risk – identification of undesirable event	Risk code	Risk assessment			Current safety precaution	Control method
			C	*L*	*RF*		
IS administration	Creation of uncontrolled back-up of date and database during IS implementation	1.01	32	10	320	No	No
	The back-up media loss	1.02	16	5	80	Storing back-up media to the safe	Documentation for data administration procedure
IS operating	Fire (over heat)	1.03	8	5	40	No	No

Object risk factor average: **147**

Risk assessment summary (risk class): **Risk class 2** that means increased of risk level, the increase of security level, standards and monitoring of security process is recommended.

Precaution proposal

Risk code	Proposal	Implementation description	Responsible person	Deadline	Estimated cost	Really implementation date
1.01	Procedures and use of model data during implementations IS	IT manager has to prepare procedure for IS implementation	IT manager	15.9.2005	50 EUR	

Table 3
Risk analysis form – example of object risk assessment

Object risk assessment summary

Object code	Name	Object risk factor average	Risk class	Objects (assets) value* [EUR]
0.1	Information system	147	2	6 500
0.2	Physical network architecture	20	1	2 500
0.3	Users desktops	170	2	1 500
0.4	Server	250	3	3 700
0.5	Patient card index	35	2	10 750
0.6	Archive	130	2	17 000
0.7	Server room	120	2	500
0.8	Work room	50	2	12 000
0.9	Laboratory	15	1	7 500
0.10	Medical office	10	1	6 800
	Average (risk factor per organization)	**109**	**2**	**88 750 EUR (sum)**

* Value are determined from accounting system as investment and overhead cost.

Precaution proposal summary

Risk code	Proposal	Implementation description	Responsible person	Deadline	Estimated cost	Priority	Really implementation date
1.01	Procedures and use of model data during implementation IS	IT manager has to prepare procedure for IS implementation	IT manager	15.9.2005	50 EUR	Yes	
1.03	Fire safety system	Installation of electronic fire signalization	Build administrator	31.12.2005	1 200 EUR	No	

Risk consequence. It is determined by Table 4. This table makes to review of an impact. The consequence is calculated as quadrate of risk level (this is expression of the risk intensity in accordance with impact and level). And the table uses the description and characterization of impact to the healthcare organization. This table very easy evaluated the consequence risk.

Table 4
Impact review

Level	Consequence	Characterization	Impact description
1	1	Insignificant	Insignificant infringement of operating procedure in the organization with immediate correction, none financial loss.
2	4	Minor	Low financial loss, infringement of operating discipline, infringement of operating instructions in the IT/SI area.
3	8	Moderate	Unauthorized operating with information in the organization, infringement of security policy and legislative rules, major financial loss, decrease of patients.
4	16	Major	Damage to the goodwill, major decrease of patients, major financial loss, escape insignificant information outside organization, wilful infringement of operating discipline, uncontrolled and unauthorized operating with information and data outside organization.
5	32	Catastrophic	Enormous financial loss, heavy decrease of patients, escapes sensitive information and personal data about patients outside organization, permanent damage to the goodwill.

Likelihood measures the chance or probability of an outcome or event occurring within a finite universe of possibilities or time. We can estimate the likelihood on experience and knowledge basis or from records of incidents and events. The Table 5 shows the likelihood of incident or event in fifth difference level. The different levels reflect the weight of likelihood for the risk factor calculation.

Table 5
Risk likelihood

Risk likelihood	Weight
Practically impossible incident or event.	1
Uncommon incident or event, but coming in specific situation.	5
Possible incident or event, incident or event was identified before.	10
Frequent incident or event.	15
Very often recurrent event and incident.	20

Risk factor (RF) is the numeric value of a risk. Risk factor is the product of likelihood (L) and consequence (C) calculated by Eq. (1):

$$RF = L \cdot C. \tag{1}$$

4.3. Result of risk analysis

Determination of risk class is needed for the improvement of process and ISMS. The classification the risk to the classes gives recommendation how to improve and take safety precaution. The risk factor can be used as classification criterion for risk categorization. In accordance with criterion we can determine risk class for improvement of processes and ISMS. The risk classes are defined by Table 6. The risk class is recorded in the risk analysis forms.

Table 6
Risk class matrix

Risk class	Description
1 ($RF \leq 20$)	Standard risk, the respect of security policy and standards procedures is recommended.
2 ($20 < RF \leq 240$)	Increased of risk level, the increase of security level, standards and monitoring of security process is recommended.
3 ($240 > RF$)	High level of risk, the acceptance of technical, organization, personal and physical proposal is recommended. The organization should be monitoring the security process and re-evaluate the security policy and standards.

Determination of object risk factor and organization risk factor is the next step for summary of process result analysis. The object risk factor can be calculated from all risk factors which were calculated for processes in the object. Average of all objects risk factor determines total risk factor the healthcare organization. The object and total risk factor can be categorized to the risk classes in accordance with Table 6 too.

4.4. Precaution proposal

The precaution proposal is the part of risk analysis for risk elimination and improvement of security processes. The proposal should be implemented if the risk factor exceeds rated value. The rated value of the healthcare organization is for example 180. The limit can be set individually for other type organization. All precaution proposals are summarized to the action plan. For other object can be identified identical processes of course. But the risk and risk factor of process can be different in dependence on processing information in the object.

5. Conclusion

Real subproject outputs are:

– audit checklists for audit of legal requirements under the GoodPriv@cy certification;

– methodology of risks analysis and its implementation at the healthcare organization;

– GoodPriv@cy certification.

The risk analysis is simple methodology developed for risk assessment and risk management as part of implementation of ISMS in the healthcare organization. Risk analysis methodology, which was design by team, was implemented in the company as part of quality management too. The team cooperated with IT administrator and responsible managers during subproject solution. Cooperation with healthcare organization on the DMS building was very interesting experience from the healthcare area for project team.
The example of the risk analysis is shown in the risk analysis forms (see Tables 2 and 3). The results of risk analysis will be used for design of action plan. The goals of action plan are elimination of risks with high risk factor.
The elements used in accordance with ISO/IEC 17799 [6], which were applied for the clients (patients) information system and databases and for the GoodPriv@cy certificating, are written in this paper in Section 3.
The healthcare organization has the security project in accordance with the legislative requirements too. This project summarizes basic ISMS policy, security analysis and contains physical, technological, personal and organizational proposal, security procedures and incident control, defines the requirements documentation and control of documents, etc. The security project and risk analysis methodology were implemented to the quality management too.

References

[1] *Information Security Management Systems – Specification with guidance use*, BS 7799-2:2002. London: British Standards Institution, 2002.

[2] F. Steiner, J. Tupa, and V. Skocil, "Risk management in manufacturing enterprises", in *Conf. ICPR-18*, Salerno, Italy, 2005, pp. 1–5.

[3] W. Ozier, "Introduction to information security and risk management", http://www.theiia.org/itaudit/index.cfm?fuseaction=forum&fid=543

[4] "Inftroduction to risk analysis", http://www.security-risk-analysis.com/introduction.htm

[5] J. Tupa, "Data Management System Audit according to GoodPriv@cy", DIGI Q – SMEs Project Works. Prague: Czech Association for Quality Certification, 2004.

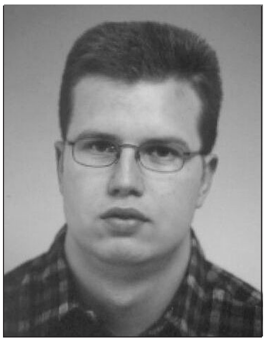
[6] *Information technology – Security techniques – Code of practice for information security management*, ISO/IEC 17799:2005 ed. 2.

[7] *Quality management systems – Fundamentals and vocabulary*, ISO 9000:2005 ed. 3.

[8] *Information technology – Guidelines for the management of IT Security – Part 3: Techniques for the management of IT Security*, ISO/IEC TR 13335-3:1998 ed. 1.

[9] *Information technology – Guidelines for the management of IT Security – Part 4: Selection of safeguards*, ISO/IEC TR 13335-4:2000 ed. 1.

[10] *Information technology – Guidelines for the management of IT Security – Part 5: Management guidance on network security*, ISO/IEC TR 13335-5:2001 ed. 1.



Jiří Tupa was born in Czech Republic. He received the M.Sc. and Ph.D. degrees in electrical engineering from Faculty of Electrical Engineering, University of West Bohemia in Pilsen in Czech Republic, in 2002 and 2006, respectively. The main his research interests are business management for electrical engineering industry, integrated management system, electronics technology, information and copyright law, information security. He is lecturer at Department of Technologies and Measurement University of West Bohemia in Pilsen.
e-mail: tupa@ket.zcu.cz
University of West Bohemia in Pilsen
Faculty of Electrical Engineering
Department of Technologies and Measurement
Univerzitní 26
CZ 306 14 Pilsen, Czech Republic

František Steiner was born in Czech Republic. He received the M.Sc. and Ph.D. degrees in electronics from Faculty of Electrical Engineering, University of West Bohemia in Pilsen in Czech Republic, in 1996 and 2001, respectively. The main his research interests are information systems, information security and technology of electronics. He is lecturer at Department of Technologies and Measurement University of West Bohemia in Pilsen.
e-mail: steiner@ket.zcu.cz
University of West Bohemia in Pilsen
Faculty of Electrical Engineering
Department of Technologies and Measurement
Univerzitní 26
CZ 306 14 Pilsen, Czech Republic

Paper

A roadmap towards healthcare information systems interoperability in Greece

Alexander Berler, Anastassios Tagaris, Pantelis Angelidis, and Dimitris Koutsouris

Abstract— The advantages of the introduction of information and communication technologies (ICT) in the complex healthcare sector are already well known and well stated in the past. It is common knowledge that in order to install any type of information system in healthcare, six main groups of issues have to be dealt with: organizational and cultural matters related to healthcare, technological gap between healthcare professionals and information science experts, legal requirements on the confidentiality of personal data, of patient related data and on data privacy, industrial and market position of healthcare informatics and interoperability complexity, lack of vision and leadership of the health care managers and health authorities and user acceptability and usability of the proposed information systems. In order to meet these issues stated above, a special focus group (Z3) performed an assessment of the situation of healthcare informatics in Greece and of the main key points that would lead to success. In that sense it is now common knowledge that Greece is lagging information and communication technology progress in healthcare because almost none of the above mentioned issues were dealt with. This assessment is the result of the interaction of more than 150 decision makers, medical informaticians, healthcare practitioners and other individual involved in healthcare. As a conclusion, this focus group resulted in 4 major propositions that will lead to healthcare informatics introduction with better success chances: focus on terminologies and standards, focus on interoperability and information systems sustainability, focus on clear goals and system metrics that can create a healthcare performance management cockpit, and focus on people and what they have to say, by creating a e-health forum. These conclusions were taken into consideration by the Greek government and are incorporated the IASYS project, the national healthcare informatics framework for the next ten years.

Keywords— interoperability, HL7, regional healthcare information system, medical informatics standards, medical codifications, information and communication technologies developments in the healthcare.

1. Introduction

The advantages of the introduction of information and communication technologies (ICT) in the complex healthcare sector have already been depicted and analysed in the healthcare informatics bibliography [1–6]. It is nevertheless paradoxical that, although several major technological discoveries such as magnetic resonance imaging, nuclear medicine and digital radiology, which facilitate improvement in patient care, have been satisfactorily embraced by the medical community, this has not been the case with healthcare informatics. Thus, it can be argued that issues such as data management, data modelling and knowledge management have a long way to go before reaching the maturity level that other technologies have achieved in the medical sector.

A variety of reasons could be proposed for this issue, though with a short analysis it becomes rather clear that modern ICT present integration problems within the healthcare sector because of the way the latter is organised. Healthcare is a strongly people-centred sector in which ICT has been considered more as an intruder, as a "spy" to the healthcare professionals' way of doing things and as a competitor to this people-centred model. Thus, if ICT intend to prove its advantages towards establishing an information society, or even more a knowledge society, it has to focus on providing service-oriented solutions. In other words, it has to focus on people and this has not been the case in most of the circumstances. It is common knowledge that in order to install any type of information system in healthcare, six main groups of issues have to be dealt with [7, 8]:

1. The organizational and cultural matters related to healthcare. This issue is rather important, regardless of any information system, since organizational models and culture do endorse neither the continuity of care, nor any type of structured data collection. Issues such as mistrust between different specialists, between the different healthcare structures or between doctors and nurses prevent in many cases the effective sharing of information. Health reforms are currently under way in many countries stressing the will to deal with this problem.

2. The technological gap between healthcare professionals and information science experts. Doctors are often reluctant to use information systems which they believe are not designed for them. From another point of view, healthcare informatics have been introduced in healthcare institutions mostly on pilot-based projects aiming at addressing specific issues and have proposed solutions addressing a small number of healthcare practitioners, resulting in establishing a complex map of information niches. This approach is the consequence of applying information technology to procedures that where not designed for it, thus creating a panspermia of information models

which are neither compatible nor interoperable, even within a single institution's environment. Efforts in creating interoperability standards and protocols, such as health level seven (HL7), are proposing solutions to address this issue, thus enabling data manipulation and knowledge management.

3. The legal requirements on the confidentiality of personal and patient related data and on data privacy. It is clear that if this issue is not addressed at a managerial and procedural level by imposing suitable policies to meet these requirements, there is little chance that medical data will be kept digitally in a structured manner (thus allowing the transition from digital islands of clinical data towards a structured electronic healthcare record). The implementation of an information system, where the electronic healthcare record is considered to be the core of the system (patient-centred model), is the only way to drive data management towards creating new knowledge. The complexity of the problem can be explained if one just observes the course of implementation of both the *Health Information Privacy and Accountability Act* (HIPAA) in the US and *Directive 95/46/EC* in the European Union (EU). The issues seem to have been dealt with at the strategic level, but still a lot has to be done in the implementation and setup of those strategies.

4. The industrial and market position of healthcare informatics. In general the healthcare market is seen by the industry as large in size but not highly profitable, mainly due to the lack of standards in implementing and interoperating healthcare informatics products. As a consequence, the industry has focused on creating mostly small scale products (i.e., laboratory information systems – LIS, radiology information systems – RIS, clinical information systems) and not on evangelising the production of information systems that are dealing with healthcare as a whole. The lack of end-to-end solutions is dealt with by interconnecting heterogeneous information systems (a rather complex task with constant change management issues) and by introducing solutions from other business sectors (i.e., ERP, SCM, CRM) that have often been rejected by "key users" as non compliant with their job description. Nevertheless, the new web technology approaches (web services, XML, etc.) and the new information technology strategies (i.e., service oriented architecture) could be the drivers towards merging information technology and healthcare services and thus enabling the establishment of service oriented products.

5. The lack of vision and leadership of healthcare managers and health authorities, and the lack of willingness to re-engineer healthcare processes for the benefits of efficiency and quality of care delivery. Some countries are in the process of introducing or implementing such business process reengineering projects in order to address healthcare delivery in a more information flow conformant way. This is a key point in reaching knowledge management, knowledge reuse and sharing, and finally proposing a solution for the knowledge-based society of tomorrow. This issue should be dealt with by proposing strategies that focus on processes and by establishing key performance indicators, balanced scorecards, or other metrics that are the upper level of a structured information flow-based model.

6. User acceptability and usability of the proposed information systems. This issue is the one most strongly related to the problem of dealing with the people-centred approach of the healthcare sector. This issue deals with information systems' user friendliness and attractiveness, with usability issues such as the time to reach a data entry point, the speed of information retrieval, the quality of information retrieval, the complex security procedures, etc. In order to implement information systems and knowledge management systems, education and training must be addressed with high priority since user acceptability is strongly related to them. Service oriented models and patient-centred information systems have a higher chance of passing the user acceptability test. A system that is not accepted by the user is often a system with poor data quality (or no data at all) and knowledge management, business intelligence or data warehousing solutions are consequently inoperable and unsuccessful.

Taking the above issues in mind, as well as the ongoing efforts of the Greek Ministry of Health, the Greek e-business forum[1] initiated a new focus group regarding e-health and interoperability, which took the codename Z3. This focus group gathered more than 150 decision makers, medical informaticians, healthcare practitioners and other individuals involved in healthcare. The focus group started working in September 2004 and ended in April 2005 with a one day event (workshop) to present publicly its findings and recommendations. The following paragraphs are depicting the result of that effort in Greece.

2. Defining the open issues

The focus group prepared an exhaustive questionnaire that was filled by the focus group members. The following list of issues was depicted from those questionnaires:

1. Political issues are strongly biasing the government's decision making strategy. In that sense, politics tend to change continuously, creating a lack of high level strategy.

[1] See http://www.ebusinessforum.gr

2. There is no national strategy for medical terminology, information systems security, disaster recovery, data interchange protocols, etc.

3. Greek medical institutions are understaffed regarding their need for the successful adaptation to new information and communication technologies.

4. As the public sector is concerned, the Focus Group noticed that procedures do not comply to the introduction of ICT, thus creating a draw-backing inertia of the national healthcare system.

5. High level leadership mostly focuses on day to day management than towards introducing the necessary structural changes to support ICT.

6. There is a strong lack of vision amongst leadership, starting top down from the high level administration.

7. The Greek medical ICT market is very small to enforce correct bottom up solutions, thus existing solutions simply follow the complex and bureaucratic way of doing things in the Greek public medical institutions.

8. The user requirements and technical specifications proposed to the implementers often lack of severity, clarity and business scope.

9. There is no follow up of other worldwide best practices, and visionaries are restricted to deploy strategies that never succeeded to overpass the design phase.

10. The proposed time management of the government ICT project is unrealistic and does not take into consideration the complexity of the healthcare sector.

11. Fund management and human resources management is not clear and are both mostly spent in unrealistic projects that to not promote ICT as success cases.

12. The high level leadership lacks of ICT knowledge and cannot focus correctly upon the benefit of the correct introduction of integrated information systems in Greek medical institutions. A large majority of questionnaire reported a techno phobic approach of the political and administrative leadership.

13. The Greek healthcare sector has four decision making groups (Ministry of Health, Ministry of Education, Ministry of Social Welfare and Ministry of Defence), thus making the business rules extremely bureaucratic, creating a business environment that lacks of homogeneity in matters of terminology and procedures.

14. The social security sector is also extremely complex and not homogenised in procedures, insurance coverage, and support to citizens. This is due to the separate route that each ministry has followed for its institutions. Even today with the operation of a general secretariat for social security, the Greek government has not succeeded yet to create the correct environment for the citizen, despite the efforts of the last years.

15. The human factor lacks of expertise and training in ICT, thus making almost impossible to locate the correct amount of key users or early adopters to promote ICT.

16. It is extremely difficult to implement business reengineering projects in the public sector. Nevertheless, many efforts are in the process of implementation.

17. The reaction to change is quite large, since techno phobia has passed from top management to a large number of employees, thus creating a hostile environment for ICT visionaries.

3. Interoperability roadmap prerequisites

In order to establish an interoperability roadmap, three prerequisites have to be met:

– slection of an interoperability architecture;

– pilot testing to establish possible open issues and implementation risks;

– defining an information systems sustainability scorecard.

3.1. Proposing an interoperability architecture

In 2001 a reform of the Greek national healthcare system [9] was introduced in order to enhance the performance and control of healthcare provision in Greece. One of the main changes was the division of the country in 17 autonomous healthcare regions where the regional healthcare authorities (RHA) are responsible for the regional healthcare strategy. In order to support this reform a series of ICT oriented interventions were introduced. After a period of analysis and design the Greek government started issuing a number of extremely detailed (more than 500 paged each) request for proposals (RFP) for each RHA [10].
The integration of existing and forthcoming information systems represents one of the most urgent priorities in order to meet the increasing clinical, organizational and managerial needs [11, 12]. In that context, the use of standards is essential since data processing needs vary widely in the complex regional healthcare environment. All RHA have a major concern in evaluating the existing operational hospital information systems and other information system infrastructure in order make a decision on whether to maintain or replace them. In Greece, more than ten distinct vendors have installed healthcare IT related products (hospital information system, laboratory information system, radiology information system, etc.) that mostly work indepen-

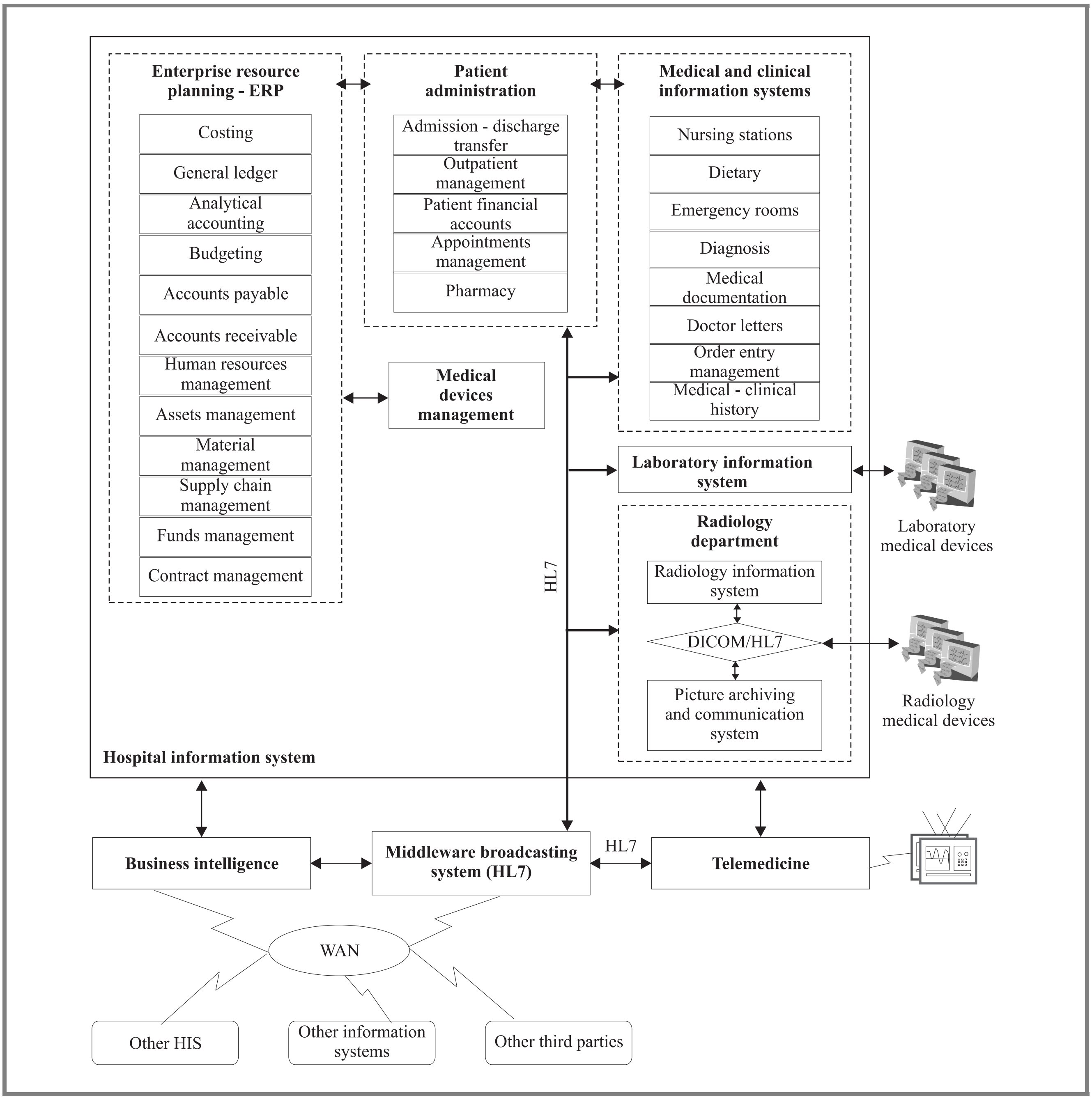


Fig. 1. Regional healthcare information system basic software components.

dently as IT niches. It is known that the lack of healthcare information standards is one barrier to the broad application of IT in health care units. The inability to share information across systems and between care organizations is just one of the major impediments in the health care business's progress toward efficiency and cost-effectiveness, as well as, the absence of a unique national or even regional patient identifier in Greece. Integration of these existing diverse systems with the future information systems to come remains problematic with a number of competing approaches, none of which alone represents the perfect solution. Current practice shows that the most promising approach to achieve a regional healthcare information system is to use, where applicable, an HL7 message-based communication system implemented by an asynchronous common communication infrastructure between healthcare sites.

The proposed information system in the RFP consists of a series of subsystems as depicted in Fig. 1, covering information management issues in a regional healthcare system. The system is innovative in the sense that it required the design and implementation of a complete and integrated information system at a regional level that comprises all types of healthcare levels (primary care, secondary care, home care, etc.), that includes interoperability issues, that covers most of the needed components and that could be able to work efficiently in a secure wide area network (i.e., a VPN) to ensure data privacy and confidentiality.

Through the aforementioned RFPs, the need has arisen to make healthcare information systems in Greece to work together as the components of regional healthcare network (RHN), where newly introduced information systems must communicate with systems already present in various healthcare institutions. The proposed solution features the use of middleware broadcasting systems that are based on information exchange via messages utilizing some application protocol (ISO-OSI level 7).

The proposed architecture fulfils at least the following requirements:

1. Existing systems do not need to be altered.

2. No significant extra (hence unanticipated) load on existing systems is introduced.

3. Connecting existing systems is an economical viable activity.

The three requirements are met by an *asynchronous message* based information exchange infrastructure defining a uniform interface for any system that must send or receive information. All systems are connected, through a uniform interface, to an interoperability framework or more technically to a common communication infrastructure (CCI). In an asynchronous message based CCI, information is exchanged between two systems by breaking up the information into chunks. These "chunks" are called *application protocol data units* (APDU). An APDU has an explicit structure that is defined by the APDU (or message) *syntax*. Additional encoding and decoding rules help sending and receiving systems to construct and to analyze APDUs. Sending systems can insert information into APDUs and receiving systems can extract information from the APDUs.

APDUs are not transmitted directly; they are embedded in so called *protocol data units* (PDU). APDUs form the "payload" of PDUs. PDUs contain enough information for

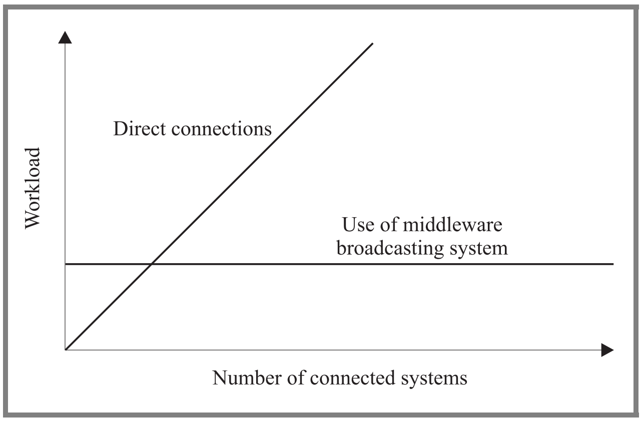


Fig. 2. Workload produced by connected systems.

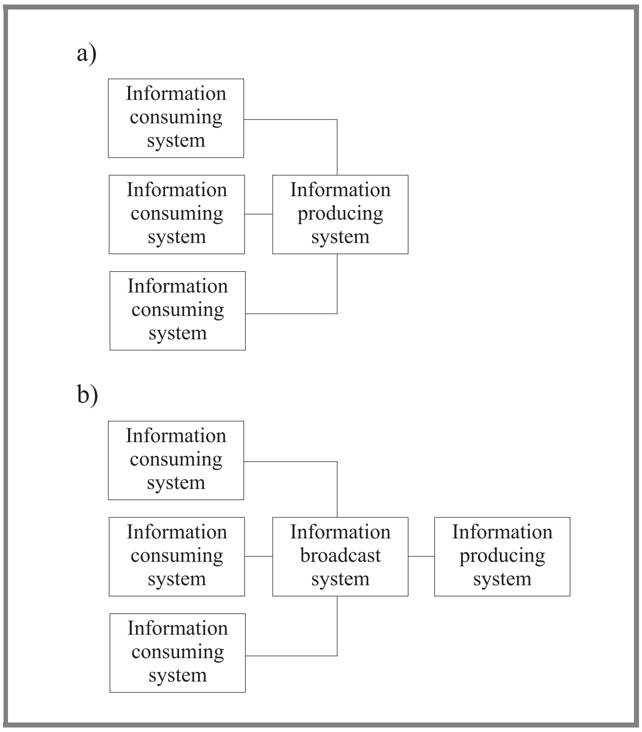


Fig. 3. Schematic representation: (a) direct connection; (b) use of middleware broadcasting system.

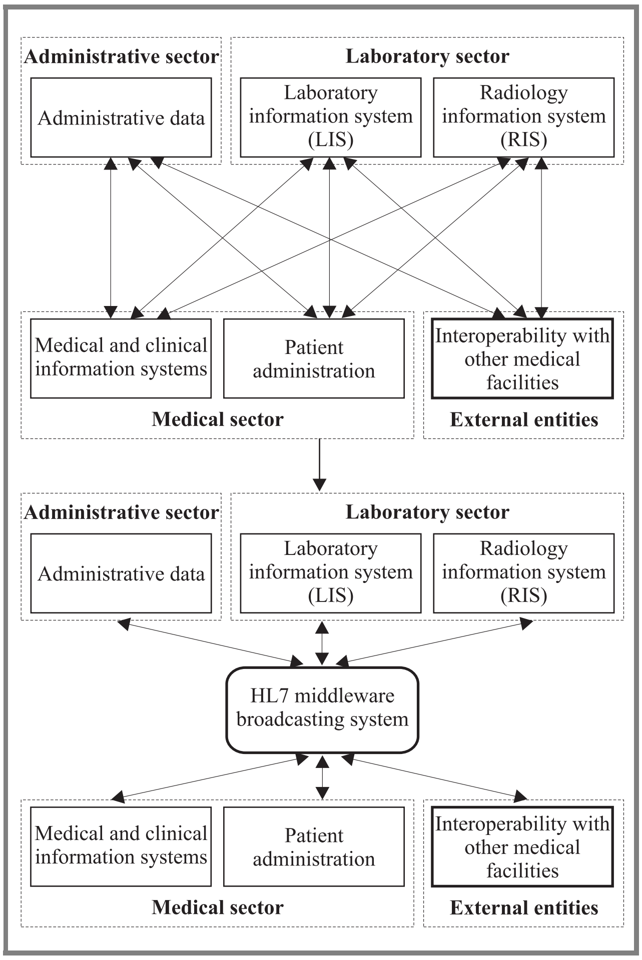


Fig. 4. Creating an interoperability framework.

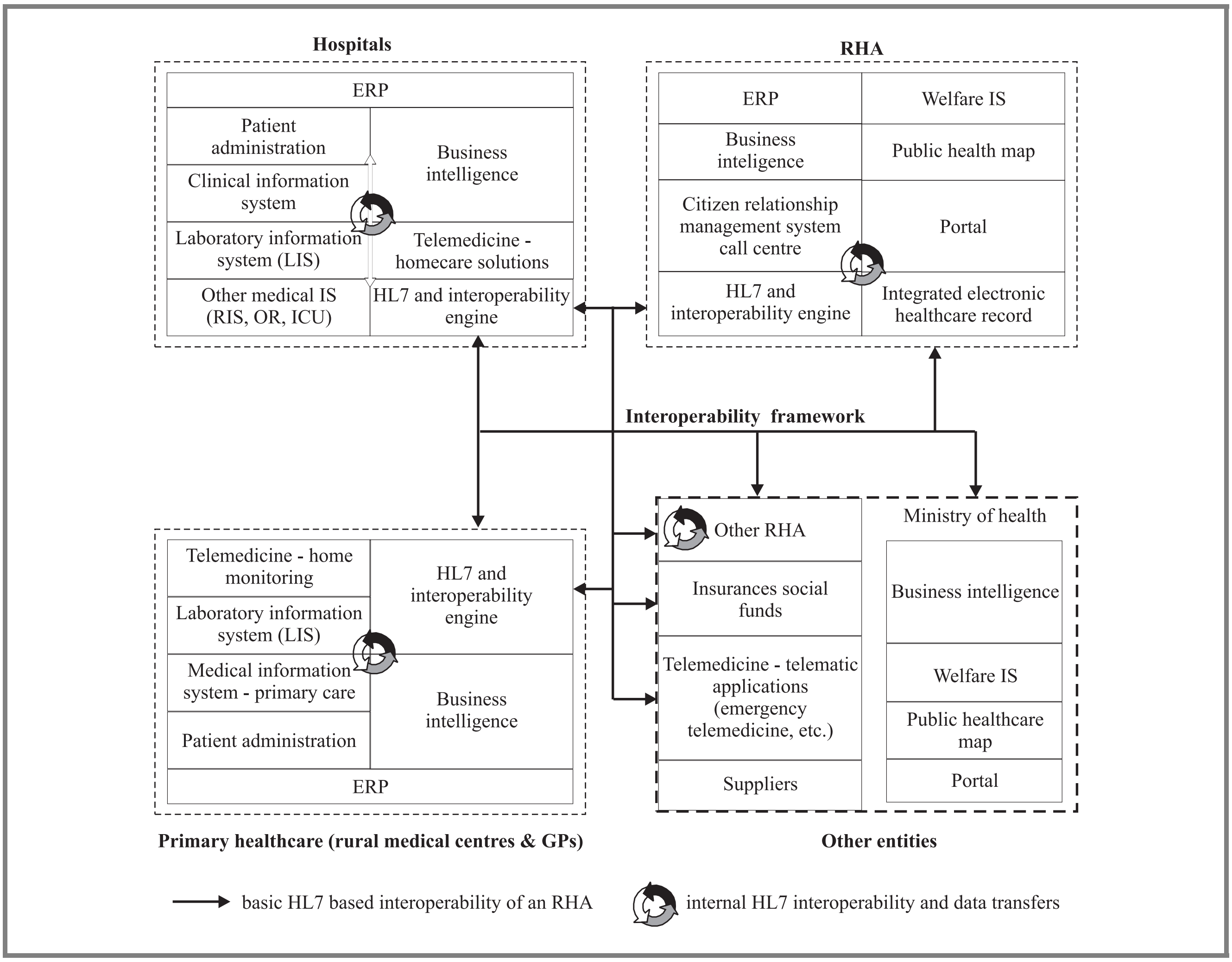


Fig. 5. Regional healthcare based interoperability framework.

the CCI to be able to "route" the information sent to the receiving application. Additional "meta" data help the receiving side to understand if the PDU has been received intact and contains the APDU anticipated.
Using (A)PDUs to exchange information between systems brings a number of distinct advantages:

1. All systems can be interfaced in a uniform way with each other.
2. There is decoupling between systems which allows information to be routed, stored and forwarded, and processed independently from the actual exchange.
3. Information exchanging does not need to reveal their internal structure to each other. This form of "information hiding" significantly improves the connectability of systems.

As depicted in Fig. 2, the use of a middleware broadcasting system is enabling the interconnection of information systems without creating extra workloads on existing information systems. When a system provides a uniform interface for sending and receiving information it can be connected easily and even routing of information becomes feasible. The latter is very important to connect remote systems that cannot communicate directly. Clearly the third advantage is the most important. The fact that two information systems do not need to know each others database schemata or database connection technology, tremendously simplifies the task of interfacing these systems. Figure 3 depicts the change that occurs when introducing a middleware broadcasting system.
Another important feature of the proposed solution is that it creates an interoperability framework that can be replicated from one healthcare institution to another. In that sense, common interoperability messages can be used to interconnect heterogeneous information systems within a healthcare institution or even at a regional healthcare level if a centralized information system is in place, as depicted in Fig. 4.
The proposed interoperability framework greatly simplifies the data exchange issue in a regional healthcare information system since a lot less interoperability connections are required and messages used are homogenized between all involved healthcare institutions.

Health level seven [13–15] is by far the most widely used message based information exchange standard in the clinical environment. It is in use on all continents of the world. Also HL7 is clearly the most mature message based information exchange standard. As a consequence, HL7 was set as a mandatory requirement in the selection process for the implementation of the RHN for each RHA in Greece.

Figures 3 and 4 mostly deal with interoperability issues within a healthcare institution, where typically hospitals are mostly concerned since they produce the wider range of medical data. Figure 5 though is extending and describing the proposed interoperability framework and clearly depict the basic interoperability paths required at a regional healthcare level. As stated before, a regional healthcare system can be either an aggregation of interconnected distributed and variable information systems, either a totally centralized system based upon an application system provider's (ASP model) approach or a combination of the aforementioned architecture. In all cases information flows, patient journey data, electronic healthcare record data (data collected from various institutions, in various formats and appointed to each individual based upon a mater patient index – MPI) are creating very important interoperability issues. It is without saying that data privacy issues are important when transferring or gathering data at a regional level and should be dealt with according to EU directives and additional national laws. Data privacy issues are addressed by the means of creating the proper patient consent mechanism, by creating and imposing strict and firm data manipulation and data storage procedures and by avoiding aggregation of sensible data when not strictly required.

Figure 5 depicts the interoperability point within a regional healthcare information system. The interoperability framework can be implemented either centrally with one middleware broadcasting system that interconnects all concerned information systems in a regional healthcare network (VPN based) setting or with an aggregation of interconnected and networked middleware broadcasting systems (one for each institution in the regional setting) that all communicate by an agreed numbers of HL7 based messages. In that sense, the cooperation of such middleware systems could be expanded nationwide, thus enabling patient mobility and data consistency within a nationwide electronic healthcare record.

3.2. Pilot testing

In order to test the proposed framework a small scale pilot project was conceived [16]. The pilot aimed at implementing interoperability among hospital information systems and the management information system (MIS) of a RHA. The implementation of an HL7 link requires fewer resources when HL7 middleware is deployed, and the data are stored in open architecture database management systems. The required human resources were one software engineer with knowledge of HL7, database systems and HL7 middleware concepts, one project coordinator (part time employment), one system administrator per site/link (part time employment), one project manager from the RHA (part time employment) and one application/database administrator (part time employment).

The time consumed for the pilot project is described in Table 1.

Table 1
Duration of the phases of such project

Activity	Duration in hours
Defining the scenario	Depends on user requirements
Defining the events and the event data	60
Selecting the right message types	24
Extend selected message types (optional)	12
Define the protocol	12
Determine implementation parameters	12
Map message fields onto table columns	40
Mapping table columns onto message fields	40
Implement message sending	120
Implement incoming message processing	120
Verification and validation	60
Total	**500 hours**

Furthermore, the risk analysis for the implementation of such projects is summarized into three categories: ***data quality, technical and organizational***. Table 2 presents an analysis of the risks.

3.3. Information systems sustainability scorecard

Defining existing information systems sustainability is not an easy task since most of the reasons for disinvesting or reinvesting in information systems is highly subjective, mostly based upon criteria such as user friendliness, cost effectiveness, etc.

The proposed scorecard is based upon some initial assumptions:

- There is no issue of sustainability concerning existing information systems that are to be replaced by technologically more advanced platforms.

Table 2
Risk analysis table

Risk	Description	Encountered during this pilot
Data quality risks		
Annotated database schema missing	There is no annotated database schema at all. Lack of a comprehensive database schema significantly reduces the likelihood of successfully reverse engineering the database schema.	No
Database schema contents not interpretable	The names used in the schema are unclear or ambiguous; also relations are not clearly defined.	No
Database tables not used as described in the database schema	Database table usage has drifted away from the logical or semantic design.	Yes
Free text fields used for structured data in an ad hoc fashion	Free text fields are using to store structured data. The database schema should be redesigned.	No
Inconsistent use of enumerable data types	In particular in reference tables like COUNTRY, CITY values like "GREECE", "Greece", "greece" all representing the country Greece.	Yes
Required data not present	Data that according to the database schema are required (NOT NULL) but are null in the tables. The database schema does not reflect the current structure of the database.	No
Semantic analysis of data	Data fields does not contain semantically valid data.	Yes
Character set encoding problems	There are problems with the character set encoding in the database.	No
Technical risks		
General interface to access database system	Access to information systems' databases through general interfaces that are independent from applications, must be available (and configured) in every (hospital) information system.	No
Proprietary operating system	The database system runs on a platform with proprietary operating system (e.g., Not Windows, Linux, Unix, or VMS).	No
Database not accessible	It is impossible to logon to the database.	No
Exotic communication protocol	The platform on which the database runs can only be connected through a non-TCP/IP communication protocol.	No
LAN not reachable	The LAN on which the pilot system is connected to cannot be reached from outside the LAN due to security matters and other reason.	Yes
Unstable or failing computing environment	The computing environment is unstable or failing causing the pilot system to malfunction.	No
Organizational risks		
Not enough competent staff	IT staff is not qualified or inadequate in quantity.	Yes
Lack of individual co-operation	Individuals in the organization are reluctant to co-operate.	No
Rules and procedures	Rules and procedures are becoming an obstacle or slow down progress especially in public services.	Yes
Lack of decision making	There are no decision makers that can put the project in progress.	No
Lack of software vender support	Software vendors that need to assist do not do this.	No
Software vendor sabotage	Software vendors are actually sabotaging the project out of commercial interests.	No

- A RHIS is an integration of specific oriented building blocks that are commonly acknowledged and agreed. Those are the enterprise resource planning (ERP), the HIS, the LIS, the RIS, the human resources management, the document management system, the interoperability middleware tool, the portal and the business intelligence tools.

As a consequence the following steps are required:

- defining the building blocks (BB) of RHIS (as above);
- proposing a scorecard for sustainability;
- defining the needs for interoperability between building blocks.

As a matter of scorecard the following criteria were proposed:

1. Technical and logical architectural conformity.
2. Recent technological platform (Windows or Unix GUI, Web GUI).
3. Interoperability capability with other BB (HL7, XML).
4. 80% coverage of ERP standard functionality.
5. 60% coverage of HIS standard functionality.
6. 100% coverage of HRM standard functionality.
7. 100% coverage of LIS or RIS standard functionality.
8. 80% coverage of the established Greek national healthcare systems procedures.
9. 100% coverage of required data exchange within BB modalities.
10. Vendor sustainability (ability to deliver and support the information systems for at least 3 years).
11. Fixed budgeting rules.

4. Medical terminology: a prerequisite for interoperability

4.1. The importance of codification

Healthcare institutions are creating a huge amount of data of any type (administrative, financial, medical, etc.) or format (reports, medical records, medical images, transcriptions, doctor letters, etc.), on a daily basis.

Despite the technological efforts and new proposed technologies of our times, a great deal of those data is still hand written or paper based, thus not enabling the exploitation of those rich information sources. Part of this delay is due to a lack of codification, terminology and standard usage for recording, storing and interchanging data. The use of medical terminologies allows systemic and procedural reuse of information in order to assist medical staff, to fill the electronic patient record, to promote prompt and correct diagnosis and to enhance quality of care.

Furthermore coded data are more malleable concerning statistical analysis and public health monitoring, both at a national and international level. Both administrative management and medical staff are able to gather any type of data fitting their job descriptions.

The most common codifications are the classifications such as the International Classification of Diseases (ICD)[2] proposed by the World Health Organization (WHO), and the nomenclatures such as Systematized Nomenclature of Human and Veterinary Medicine (SNOMED)[3]. Other types of codifications are the thesauruses, the taxonomies and the formal terminologies. In Table 3 some of the most commonly proposed and used codifications are listed more as examples than a complete list. The oldest classification reported, the "London Bills of Mortality" was conceived in England for forensic purposes in 1662. WHO started ICD in 1901 with Version 1 and today we have reached Version 10, finalized in 1992. SNOMED started in 1928 (SNOMED RT) and is been continuously updated, now having more than 361 800 medical terms, 975 000 descriptions and 1 470 000 semantic correlations in SNOMED CT (2004).

Organizations as WHO, College of American Pathologists (CAP), Health Level Seven and the world organization of national colleges, academies and academic associations of family physicians and general practitioners (WONCA) are not the only bodies that have deployed successful coded data sets. It is rather common that national standardization bodies are either translating most commonly used international codifications or creating their own national subsets of any type and complexity.

Codifications by themselves are one of the most important steps toward public health monitoring, cost containment and better healthcare services to the citizens. Codifications are also extremely important as input or output of a healthcare information system of any range and penetration. The use of coded data results in having high quality structured data that enable better reuse of the knowledge created during the day to day process of patient treatment, thus enabling patient history keeping, diagnoses recording and better healthcare outcomes. Structured data enable statisticians and administrations to better monitor public health, disease prevention and strategic policy planning.

The use of coded data is also the cornerstone of cost analysis of a well designed healthcare system, making it possible to foresee procurement requirements, institution deployment and other important decision regarding healthcare management. It is important to state that EU has a strong

[2] See http://www.who.int/whosis/icd10/
[3] See http://www.snomed.org

Table 3
Most common existing codifications (*copyright: A. Berler*)

National and European data sets	**Financial and administrative data sets**
Master patient index	Greek GL prerequisites (PD 146/2003)
Social security number (SSN)	DRGs
National and EU statistical codification sets	NCDP (e-claiming)
Patient record: OpenEHR/HL7-RIM	ECRI
	Data interfacing protocols
Medical terminology	HL7 (Version 2.x/Version 3)
ICD 10 – WHO	HL7/CDA
SNOMED – CAP	DICOM 3.0
ICPC 2- WONCA	**Other codifications**
Diseases and procedures	ATC/DDD ή NCD (Drugs)
ICD-10-PCS/CPT	GMDN
CCI –	**Research**
CCAM	Arden syntax (HL7), OWL, GELLO (HL7),
HCPCS, CPT, OPCS-4	semantic web, GLIF, XML topic maps

focus towards structured medical and clinical data and has proposed a series of white papers, green papers, communications and directives [17–22]. Codifications are also critical for the dissemination of medical knowledge and information systems interoperability. It is not possible to design any type of interoperability roadmap without taking into consideration the strategic need at a national or European level for structured data. Initially medical terminologies, clinical classifications, medical procedures and clinical guidelines were proposed as a solution to calculate and restrict the number of medical errors or adverse drug events.

A large number of studies in the US [23, 24], Australia [25], Canada, Denmark, Italy, The Netherlands, Sweden and New Zealand, all report that a large number of adverse drugs events and medical errors have resulted in damages of the health of patients. In the UK, statistics report that about 10% of inpatients have been involved in episodes of care where wrong dose or other medication was given with minor or important consequences in patient's health status. The financial costs of those events are estimated at £3 billion only for the extra bed days. As a consequence the use of e-prescribing, bar coding and/or computer based order entry systems are of critical importance and have proven to reduce dramatically the number of medical errors.

In Italy more than 14 000 patients die every year due to medical errors whilst this number reaches each year 44 000 up to 98 000 in the US, surpassing death tolls that are accredited to traffic accidents, breast cancer, AIDS, etc. [24]. All studies state that those errors could be prevented or at least a large number of them, if medical data collected had the proper quality rate. Medical terminologies and codifications have a lot to offer in that sector:

- In 1992 at LDS Hospital of Salt Lake City (US) the establishment of an adverse drug reaction monitoring system recorded 569 cases, saved many extra bed days and $1 000 000 of the hospital's budget.
- The drug "Seldane" was approved by the Federal Drug Administration (FDA) in 1985, presented the first adverse reactions with erythromycin (cardiac arrhythmia) in 1992 and was only withdrawn in 1998 due to lack of decisive data.

The use of codified data into information systems in healthcare provides the ability to those systems to interoperate and exchange important medical knowledge in order to establish a unique electronic healthcare record (EHR) for each citizen by collecting all important data from each patient encounter with the healthcare system. EHRs cannot be created with medical codifications since they are the base for any type of semantic interoperability. This issue is not new since Florence Nightingale stated in 1893:

"*In attempting to arrive at the truth, I have applied everywhere for information, but in scarcely an instance have I been able to obtain hospital records fit for any purposes of comparison. If they could be obtained, they would enable us to decide many other questions… They would show subscribers how their money was being spent [and] what amount of good was really being done with it…*"

4.2. Medical terminology and codifications: the cornerstones of e-health

Figure 6 depicts the workflows both external and internal that have to be met within a healthcare system of any range (from a single institution to a national healthcare system). From that figure it is clear that three major structural re-

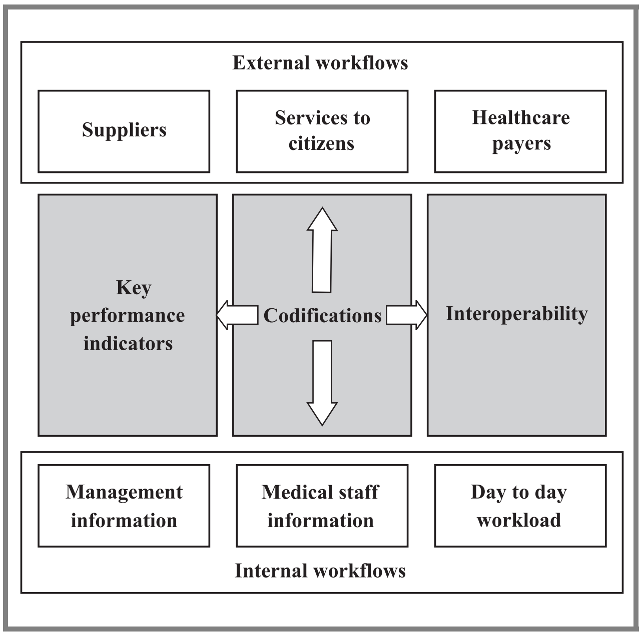


Fig. 6. Healthcare workflows (*copyright: A. Berler*).

quirements have to be met to reach a satisfactory level of coverage of the healthcare workflows with the use of ICT in healthcare:

1. The deployment of key performance indicators and public health indicators, which is one of the key recommendations and directives of the European Union [26, 27].
2. The design and implementation of an interoperability framework based upon commonly adopted and agreed international standards such as HL7 (already adopted nationally in the US, Canada, New Zealand, Australia, The Netherlands, Germany and UK, while other States such as France, Croatia, Ireland, Italy, etc., are moving towards that direction). EU is also in favour of such strategic policies [28–30].
3. The implementation and maintenance of national medical terminologies as described in the previous paragraphs.

5. Focus group suggestions

Taking all the above mentioned issues the focus group reached a consensus that is described below as a set of recommendations for the Greek medical informatics market, the establishment of medical codifications and for the establishment of an interoperability roadmap.

5.1. Ten recommendations for the Greek medical informatics market

The recommendations are as follows:

1. Urgent involvement of the leadership in favour of projects related to the introduction of ICT in healthcare.
2. Urgent involvement of key users in the design process of ICT projects instead of simple top down decision based projects.
3. Incorporation of knowledge experts such as the Greek affiliate of HL7 international.
4. Continuous ICT dissemination training programmes.
5. Top down design should be restricted to business planning and strategic objectives clarifications in order to make the national business rules crystal clear.
6. Strategic business planning continuity from the Greek Ministry of Health regardless from any political or governmental change. This should be made possible by employing a number of business experts focused towards ICT implementation in healthcare which is at least a ten years plan.
7. Strategic cooperation with other decision makers in healthcare such as the Ministry of Social Security and the Ministry of National Economy.
8. The establishment of an information authority monitored by the Ministry of Health that will be responsible for the implementation of ICT strategies, maintenance of medical terminologies and the management of the interoperability framework.
9. The creation of national public health indicators that will be in accordance to the EU guidelines and requirements.
10. Establishment of an e-health forum in order to create a constant interaction framework between all key players in the healthcare sector (government, medical institutions, industry, medical informatics implementers, etc.). This forum will be responsible for the public concertation of regulations, terminologies, strategies and other policy papers so that the maximum consensus can be reached before the implementation of new strategies and regulations.

5.2. Ten recommendations for medical codifications

The following recommendations were drawn up:

1. Greece has to fully participate in the creation of international standards and protocols by assigning national delegates to all forums and standard development organizations related to healthcare.
2. The "one size fits all" codification scheme is not efficient as medical specificities are regarded. Some clinicians prefer nomenclatures (i.e., pathologists) while others prefer simple classifications (i.e., internists).

3. There is an urgent need to select and implement the Greek set of codifications since the existing scheme of "any code is good" is a major draw back for any national data quality strategy and national ICT deployment for better health and cost containment. This should be made clear to all decision makers and administrative leadership.

4. The deployment of medical terminology could be assigned to the healthcare market itself through scientific societies and international standardization bodies.

5. EU directives should be taken into immediate consideration regarding medication related errors, the creation of access-free libraries of codes, the creation of workflow models based upon adopted standards (OpenEHR, HL7-RIM) and abstinence from the creation of national standards where international or European standards are already in place.

6. For the successful implementation and use of medical terminologies it is required that medical personnel is immediately involved in the design and proposition process, constant dissemination and training strategies are followed, consensus based decision making is adopted, job descriptions are refined and incentives are proposed.

7. Each selected codification should be selected for the specific requirements that need to be covered. All codifications should be maintainable and upgradeable and have the possibility to interrelate with other terminologies.

8. Codifications and terminologies should be selected as integrated parts of a wider interoperability platform so that all type of internal or external workflows can be completed with the use of ICT.

9. A constant dialogue framework must be established regardless of any political matters and governmental changes. The proposed e-health forum is an optimum solution for this clause. A five to ten years consensus is absolutely necessary.

10. Greece has separated the strategic planning of healthcare and social security, thus cutting the correlation of healthcare providers from payers. This has created a duplication of standardization efforts not always pointing to the same direction; Best practices in interoperability and standardization in health have often started from the payers rather than from the providers. Payers, providers and patients should be put all together under the same strategic umbrella as soon as possible to create the needed economies of scale.

5.3. Ten recommendations towards interoperability: creating the roadmap

There are the following recommendations:

1. Deployment of an interoperability framework based upon common communication interfaces.

2. Assessment and sustainability of existing information systems in medical institutions, based upon a specific scorecard methodology.

3. The healthcare informatics market should strongly focus towards standards conformance and standards maintenance. Consensus based processes for the deployment of the basic standards functionality are of critical importance (i.e., implementing integration labs).

4. HL7 is mature enough to solve most of the interoperability issues in Greek and many more than simple data interchange.

5. HL7 standards should be refined to meet peculiarities of the Greek healthcare system if such issues exist.

6. HL7 Hellas can assist the Greek ministry of health in the required standardization process that is needed to implement a national interoperability platform (terminologies, processes, workflows, performance indicators, etc.).

7. Specific task forces and standardization teams should be established immediately, under the umbrella of an information authority or of an independent scientific society, such as HL7.

8. National interoperability conformance statements must be implemented based upon the work done by integrating the healthcare enterprise (IHE) with the use of HL7 and DICOM conformance statement templates and methodologies.

9. Greece should follow the work done by international task force created by standardization bodies such as ISO, CEN/TC 251, HL7, openEHR, etc. This is especially valuable as the creation of a national EHR is regarded.

10. Immediate involvement of Greek experts and knowledge workers in international standardization processes.

6. Conclusions

The result of the focus group was publicly presented during a one day workshop with the involvement of all key players of ICT in healthcare in Greece. It was not expected that those recommendations would change the situation in Greece overnight. Nevertheless the situation of ICT introduction in Greece is blooming of activity with more that 15 high level ICT projects in the implementation process, with the initiation of consensus based efforts in order to reach a national framework regarding uniform workflows and processes, medical terminologies and an interoperability roadmap. Concerning the latter, HL7 has been largely adopted by the project implementers and the sustainability scorecards is expected to integrate information niches wherever this is plausible, both technically and financially.

Finally, the Greek government has made tremendous efforts into proposing a complete strategy regarding ICT in Greece, by proposing the IASYS project to be gradually implemented in the forthcoming years. All projects under way at this moment will be integrated into this national strategy from a technical an procedural point of view. The proposed interoperability roadmap will permit information systems to cooperate efficiently and work together as one. In order for this important strategic project to succeed an information authority is to be established to manage centrally all ICT projects in Greece and in parallel to create the required coded data structure, procedures, workflows and quality rules for these information systems. In addition, the Ministry of Health is about to announce the operation of an e-health forum covering most of the proposed recommendations of focus group Z3.

In that sense most of the focus group recommendations have been considered by the Greek Ministry of Health. The process of the successful introduction of ICT in the Greek healthcare system should nevertheless take more than ten years to be completed.

References

[1] L. Eder, *Managing Healthcare Information Systems with Web Enabled Technologies*. Hershey: Idea Group Publishing, 2000.

[2] S. Englebardt and R. Nelson, *Healthcare Informatics, an Interdisciplinary Approach*. St Louis: Mosby Edition, 2002.

[3] A. Harmoni, *Effective Healthcare Information System*. Hershey: IRM Press, 2002.

[4] T. E. Norris, S. S. Fuller, H. I. Goldberg, and P. Tarczy-Hornoch, *Informatics in Primary Care, Strategies in Information Management for the Healthcare Provider*. New York: Springer, 2002.

[5] E. H. Shortliffe, L. E. Perreault, G. Wiederhold, and L. M. Fagan, *Medical Informatics, Computer Applications, Healthcare and Biomedicine*. 2nd ed. New York: Springer, 2001.

[6] R. Stegwee and T. Spil, *Strategies for Healthcare Information Systems*. Hershey: Idea Group Publishing, 2001.

[7] I. Iakovidis, "Towards a health telematics infrastructure in the European Union", in *Information Technology Strategies from US and the European Union: Transferring Research to Practice for Healthcare Improvement*. Amsterdam: IOS Press, 2000.

[8] I. Iakovidis, "Towards personal health record: current situation, obstacles and trends in implementation of electronic healthcare records in Europe", *Int. J. Med. Inform.*, vol. 52, no. 123, pp. 105–117, 1998.

[9] N2889/2001 (FEK-A/37/02.03.2001), Greek National Healthcare System Reform Act, 2001 (in Greek).

[10] Information Society SA, "Healthcare information system for the 2nd regional healthcare authority of Central Macedonia", Request for proposal co-funded by the 3rd CSF under the EU decision C(2001)551/14-3-2001, Greece, May 2003 (in Greek).

[11] S. Spyrou, P. Bamidis, I. Chouvarda, G. Gogou, S. M. Tryfon, and N. Maglaveras, "Health care informatics standards: comparison of the approaches", *Health Inform. J.*, vol. 8, no. 1, pp. 14–19, 2002.

[12] J. Grimson, W. Grimson, and W. Hasselbring, "The SI challenge in health care", *Commun. ACM*, vol. 43, no. 6, pp. 49–55, 2000.

[13] "Health Level Seven (HL7): an application protocol for electronic data exchange in healthcare environments", Version 2.1, 1990. Chicago: Health Level Seven, 1990.

[14] W. E. Hammond, "Health Level 7: an application standard for electronic medical data exchange", *Top. Health Rec. Manag.*, no. 11, pp. 59–66, 1991.

[15] HL7 standards Internet resources, http://www.hl7.org

[16] S. Spyrou, A. Berler, and P. Bamidis, "Information system interoperability in a regional healthcare system infrastructure: a pilot study using healthcare information standards", in *Proc. MIE 2003*, Saint Malo, France, 2003, pp. 364–369.

[17] OECD, STI Scoreboard, A.6.3, Health-related R&D, 2001.

[18] CEN/ISSS e-health standardization focus group, 2004, http://www.centc251.org/ehealthfocusgroup.htm

[19] P. Doupi and P. Ruotsalainen, "eHealth in Finland: present status and future trends", *Int. J. Circ. Health*, vol. 63, no. 4, pp. 322–327, 2004.

[20] P. Doupi and P. Ruotsalainen, "Healthcare informatics in Finland: current status and future prospects", *British J. Healthc. Comput. Inform. Manag.*, vol. 21, no. 10, pp. 20–23, 2004.

[21] "MEDITRAV WP11, Deliverable 1", P. Doupi *et al.*, Eds., in *The State of eHealth in Europe*, Helsinki, Finland, 2003.

[22] M. McKee and J. Healy, "Hospital in a changing Europe", "European Observatory on Health Care Systems" Series, Buckingham – Philadelphia, Open University Press, 2002.

[23] HMSO, "An Organisation with a Memory", Report of an expert group on learning from adverse events in the NHS chaired by the Chief Medical Officer, England, 2000.

[24] L. T. Kohn, J. M. Corrigan, and M. S. Donaldson, *To Err Is Human: Building a Safer Health System, Institute of Medicine*. Washington: National Academy Press, 2000.

[25] QAHCS, "Quality in Australian Health Care Study", 1995.

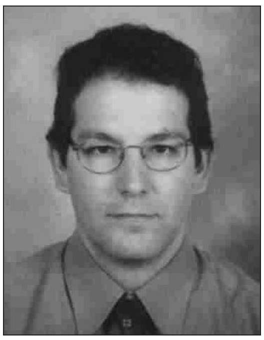
[26] F. H. Roger, "The minimum basic data set for hospital statistics in the EEC", Document EUR 7162, Luxembourg, Office of Publications of the EEC, 1981.

[27] EU Council, "Decision No 1786/2002/EC of the European Parliament and of the Council of 23 September 2002 adopting a programme of Community action in the field of public health (2003–2008)", 2002.

[28] EU Commission, "e-Europe 2005: an information society for all – an action plan", COM (2002) 263 final, 28.05.2002.

[29] EU Commission, "e-Health – making healthcare better for European citizens: an action plan for a european e-health area", COM (2004)356, 30.04.2004.

[30] EU Commission, "Follow-up to the high level reflection process on patient mobility and healthcare developments in the European Union", COM (2004) 301 final, 20.04.2004.



Alexander Berler was born in Lausanne, Switzerland, in 1969. He received his degree in electrical engineering from the Aristotle University of Thessalonica, Greece, in 1995 and his M.Sc. degree in biomedical engineering from the National Technical University of Athens (NTUA), Athens, Greece, in 1997. He is currently pursuing a Ph.D. degree at NTUA in medical informatics focusing on the design and development of interoperable healthcare information systems towards the implementation of a citizen virtual medical record. He has been with the Department of Electrical Engineering, NTUA, since 1996 as a Research Postgraduate Student working in the area of biomedical engineering in European Union funded projects and he is currently working at Information Society SA as a project manager in large healthcare informatics government project co-funded by the European Union under the 3rd Community Support Framework. He is a student member of IEEE Computer Society, IEEE EMB society since 1998, a member of ACM and European Society on Engineering in Medicine (ESEM) since 1998, and a BoD member of HL7 Hellas, the Greek international affiliate of HL7 since 2003.

e-mail: berler@biomed.ntua.gr
Biomedical Engineering Laboratory
School of Electrical and Computer Science
National Technical University of Athens (NTUA)
9 Iroon Polytechniou st
15773 Zografou Campus
Athens, Greece

Anastassios Tagaris was born in N. Figaleia, Greece, in 1968. He graduated from the National Technical University of Athens (NTUA) Department of Electrical Engineering and Computer Science and received his degree in this area in 1993. He obtained his Ph.D. degree in medical informatics from the NTUA in 1999. During 1993–1994 he was working for the Research Committee of NTUA, for the development of the financial management system of the European projects undertaken by NTUA. From 1999 he is working as a teaching and research assistant in the Biomedical Engineering Laboratory, Department of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece. From 2001 to 2005 he used to work in CCS SA in the Medical Informatics Department as head engineer of the Laboratory Information Management System (Medilab) and the HL7 Communication as well as the radiology information system development team. He is now working as a medical informatics researcher in the Institute of Communication and Computer Systems (ICCS). He has been active in a number of European and national R&D programs in the field of information and telecommunications technologies in healthcare. His current research interests include among others interoperability standards and applications of informatics and telecommunications in healthcare and medicine, Internet technology, applications of grid technologies in healthcare, knowledge databases and newpage data mining as well as business process reengineering, and telemedicine networks. He has published 6 scientific papers in international journals and conferences. His current research interests include health information systems and telemedicine, knowledge databases and data mining, application of grids to healthcare systems, image content based retrieval from databases, business process reengineering, and telemedicine networks. Doctor Tagaris has been active in a number of European and national R&D programs in the field of telematics applications in healthcare. He is a member of the Technical Chamber of Greece, the IEEE and the Greek Society of Biomedical Technology as well as an active member of the HL7 Hellas working groups.

e-mail: tassos@biomed.ntua.gr
Biomedical Engineering Laboratory
School of Electrical and Computer Science
National Technical University of Athens (NTUA)
9 Iroon Polytechniou st
15773 Zografou Campus
Athens, Greece

Pantelis A. Angelidis received his diploma in electrical engineering and his Ph.D. in telecommunications engineering from the School of Electrical Engineering, Aristotle University of Thessaloniki, Greece, in 1989 and 1993, respectively. Since May 2001 he holds the position of the Director of INA – Southeastern Europe Telecommunications and Informatics Research Institute. He has worked as a technology expert in the areas of telecommunications and biomedical engineering for the past 12 years. He has served as a project manager in over twenty research projects. He has presented more than 40 papers in international scientific conferences and has published 8 original articles in international research journals and over 10 in Greece. Doctor Angelidis is a member of the permanent consulting committee of the Technical Chamber of Greece on telecommunications, a certified consultant on telecommunications of the Technical Chamber of Greece (Central Macedonia chapter), a certified consultant on telecommunications of the Greek Courts (Thessaloniki), member of the CEN TC224/PT06 technical committee and Vice President of the HL7 Greek Chapter.

e-mail: pantelis@inatelecom.org
INA c/o HL7 Hellas
Southeastern Europe Telecommunications
and Informatics Research Institute
9th km Thermi Road
P.O. Box 60080
570 01 Thessaloniki, Greece

Dimitris Koutsouris was born in Serres, Greece, in 1955. He received his diploma in electrical engineering in 1978 (Greece), DEA in biomechanics in 1979 (France), Doctorat in genie biologie medicale (France), Doctorat d'Etat in biomedical engineering 1984 (France). Since 1986 he was research associated on the USC (Los Angeles), Rene Descartes (Paris) and Associate Professor at the School of Electrical and Computer Engineering at the National Technical University of Athens (NTUA). He is currently Professor and Head of the Biomedical Engineering Laboratory. He has published over 100 research articles and book chapters and more than 150 conference communications. He is also President of the Greek Society of Biomedical Technology. He has been principal investigator in many European and national research programs, especially in the field of telematics in healthcare.

e-mail: dkoutsou@biomed.ntua.gr
Biomedical Engineering Laboratory
School of Electrical and Computer Science
National Technical University of Athens (NTUA)
9 Iroon Polytechniou st
15773 Zografou Campus
Athens, Greece

Paper

Comparative analysis of models and platforms for the e-learning portals

Artem Boyarchuk

Abstract— The main basic functions of portal systems have to be implemented by the functionality of it platform. The technological platform of portal should be complied with the specified requirements. The basic set of services and components of e-learning portal contains five components for: services performance, user services, data management, portal adaptors, web infrastructure. The paper presents the detailed outcomes of well-known platforms analysis, discusses the advantages and disadvantages of mentioned systems and gives recommendations for customers in making a choice among these platforms.

Keywords— portal systems, services, components, technological platforms.

1. Introduction

The technological platform for development and support of system of e-learning portals is the software-hardware complex that allows establishing and maintaining the portals for different purpose and architecture and provides the accomplishment of the following functions: execution of the applications, possibilities for team-working, content management services, user administration, control and management of the productivity, knowledge management, support of the communications, personification, profiling, search, ensuring of security, standard www-access to the portal.

The technological platform of the e-learning portal should comply with the following requirement to provide the effective, dependable and flexible working:

- Conformity with international standards for open systems including standards for development, maintaining and documentation.
- Presence of advanced technological tools for integration with other applied systems and database management systems (DBMSs).
- Adaptability, tools for tuning of portal for functional requirements of specified university/institute or educational system, solution transfer technologies between the platforms.
- Scalability depending to the users' amount, stored data volumes, data exchange intensity, etc.

2. Requirements to the basic set of components and services

The technological platform, which has been selected for development of the e-learning portal, has to include the basic set of services and components. These requirements allow us to develop the portal with any architecture without delays and big financial resources due to ready set of prepared services and components.

The basic set of the services and components has to include five main elements: the component of presentation services, the component of user services; the component of information management; the component of portal adapters; the web infrastructure component [1]. Lets give the brief overview of mentioned items.

2.1. Component of presentation services

Presentation services allow users to manage the view of portal user interface and allows portal to dynamically accommodate the content presentation and navigation system to the device possibilities, which control the user access. Two different types of user interface are supported: *voice interface* (at the second stage of development) and *data interface*. The set of devices for establishment of the connect to the portal has to include desktop and portable computers with any web browsers, personal digital assistant (PDA), wireless devices – mobile phones and electronic pocket books. The portlet environment (services for personalized portal views [1]) and packing services permit user to decide what content type is "eligible" for used device because not all applications are available for all devices.

2.2. Component of user services

Component of user services helps users to find necessary information and people; to define what applications and data are needed for user; to collaborate with other users.

Personalization services (services of user personal portal) bring the content (data, applications and expert knowledge) into accord with necessities of concrete user. User may define the rules of preferable content view as part of his own *profile* for managing of this process. Afterwards the user may change the profile settings. Besides, the portal can change its settings dynamically on the basis of collected data set concerning the functioning of the applications and user's behavior (so called *click stream*). The usage of common filtering (of logical output) and rules allows the portal to use only important information for view personification and facilitation of the end-user working.

Security services (permission' control and authorization) give users the single-authorization access to all allowed content (applications, information and expert knowledge). Security services also guarantee the inaccessibility of the portal data and applications for those users who have not enough permission to access these objects.

Access services give users an access to information and applications by means of portal adapters and uniform resource locators (URLs).

Search (meta-search) **engines** compose the ideal interface for searching for the needed content at the internal and external sources simultaneously.

Publication services (informational services) support both automatic and manual publication. In case of manual publication the users with required permission may publish the content for shared working with other users. It includes office documents, materials for discussion groups, knowledge, external information (URLs, real-time news), rules, expert knowledge and transactions. At the publication new materials is included in the appropriate category of the portal *informational catalogue*.

Subscription services allow users to take out a subscription to interesting informational channel whereupon they will receive the notifications concerning new information or changing existing one at the selected topics. The subscription mailing may be carried out according to the *schedule scheme* or at *event coming.*

Delivery services (delivery service) manage the delivery of content to the subscribers (users and applications) according to the delivery politics. The content delivery may be implemented into the format, which corresponds to device (PC, pocket PC, smart-phone) or application that represent the given information.

Collaboration support services (interactive interaction services) give the tools for collaboration. With help of these tools users inside and outside the company can interoperate with each other and use content jointly. Document circulation services allow users to define:

– processes which have to be executed by these users;

– processes that have to be started in case of subscription, decision or personalization rule is true.

2.3. Component of information management

Informational management services include several items.

Informational catalogue (moderated catalogue of educational resources) is the method for searching for the information and applications accessible through the portal. The catalogue contains the links (metadata) to the content items grouped by topics. It indexes the structured and non-structured information, applications and other objects. Informational catalogue is often the determining factor for choice of the solution for development of the portal. It may be implemented as usual filesystem with the basic set of functions or as database with its rich possibilities.

Content layout manager contains the services for content systematization and grouping. Information collectors, which start after the user request or according to the schedule, "look through" the web sites and collect metadata about applications. New and updated content is transferring to the content layout manager, which takes the information into the appropriate section of informational catalogue.

Rules catalogue – rules are defined and processed by means of *rules editors*, which include to the personification, subscription and documents circulation management. The rules are stored at the rules catalogue and divided into the set of types:

– personalization rules manage the users access to the information blocks;

– content layout rules combine the interlinked data for facilitation of the work;

– decision-making rules manage the automatic decision-making, recommendations distribution and processes execution;

– subscription rules are intended for information delivery according to the schedule and to the request.

Event manager is responsible for job starting on the basis of rules, which are defined in the rules catalogue. The rules can start the actions such as report generation, e-mail delivery, operational transaction start, etc. At event coming the event manager notify the specified user or device and deliver the informational package with the help of delivery services and/or accomplish the other automatic actions.

Repository of shared information (the bank of educational objects) is the logical repository that is used for management of informational content created and published by common worked users.

Content unification manager is responsible for the integration of the application components and content (voice information and data) for delivery through the usual and wireless communication lines. This manager is carried out by personalization services.

2.4. Component of portal adapters

Portal adapters grant access to the wide range of information. To provide the interactions between applications, data and users the portal contains the adapters, which carry out the functions of interface between different informational systems and formats. Adapters may be supplied as the product for development of the portal, purchased at the specialized company or developed on one's own [3]. There are several types of adaptors to connect with different types of content.

Database (DB) adapters and files – application program interfaces (APIs) of DBs and files, which grant access to the data stored in different sources of structured information (customers' DBs, analytical information of BI and OLAP systems).

Adapters of knowledge instruments (BI) – API for access to the knowledge instruments, analytical software packages and knowledge portals BI.

Content management adapters – API for access and management of content management systems, which store non-structured commercial information (pictures, audio, video, etc.)

Adapters for application integration allow granting access to the package applications and old computer systems. This kind of adapters also includes modules for transaction management support.

2.5. Web infrastructure component

Application development tools include tools, portal components and application components, which can be used jointly with software development tools (web design instruments, rules editor and JavaBeans environment).

Administration and productivity control tools designed for portal systems administrators. It gives them possibilities for management of users and their groups, system productivity, content management (updating and synchronization) and system analysis. Administration tools should be integrated with other management instruments organizing the common framework for systems' management in the organization. Productivity control tools include caching tools, service distribution tools, and workload analysis tools.

Web applications server is the standard component of web infrastructure, which supports integration of corporative web applications [2]. The web applications server has to:

- support of several API applications and API servlets, JSP, JavaBeans, EJB, Corba, etc.;
- contain the HTTP server (Apache, Netscape or Microsoft IIS);
- provide the synchronous and asynchronous transaction management.

3. Comparative analysis of most well known portal platforms

The solutions of the most well-known and trustworthy portal development companies have been selected. For comparative analysis the following approaches were accomplished: questionnaire design, expert evaluations, experimental installation of portal platforms.

3.1. Hummingbird enterprise information portal (EIP) 5.0

Hummingbird company has developed the server of portal that provides the dependable access to the structured and non-structured information and applications [3]. Hummingbird portal is established at the entirely open architecture that allows setup and tuning it easily. Practically Hummingbird presents at the market the first *metadata portal*. Hummingbird CAP server provides the unified and secure access to the multitude heterogeneous catalogue services. CAP protocol delivers from the necessity to develop the specific code for providing of the interface with each catalogue service.

The main portal component is the Java-based platform-independent server, which delivers the set of tools and information required for user to perform his tasks. Inside organization Hummingbird portal increases the effectiveness of the work of analytics, managers and mobile users. Hummingbird EIP also consolidates the communications between organizations with the use of "business-business" transactions and e-commerce.

Advantages: EIP demonstrates the excellent characteristics for development of the corporate portals.

Disadvantages: the scalability, security, modularity, flexibility in decision-making is provided insufficiently.

3.2. X-Ware technological platform

X-Ware technological platform is the integrated package of software solutions designed for development and support of large distributed systems. Three-level architecture of technological platform allows connect up any quality of the external modules to the kernel ("shared bus"). These modules may be located at any ex-territorial area and processes the data at various formats and based at the different software-hardware platform. System kernel may be developed as group of distributed modules.

The main attached modules of X-Ware platform:

- content management module (provides content producing and updating, its formatting, editing and remote management);
- "stager" module (responsible for continuous development and updating of web resources);
- geo-targeting module (provides the ability for geographical localization of remote user and giving him information in accordance with his location);
- catalogue management module (designed for forming of the hierarchic structures of URLs for the usage at the universal search engine or geo-targeting module);
- search engines (include two modules: incremental search engine (for immediate indexing) and universal search engine (for periodical indexing).

Advantages: X-Ware-based systems have abilities for integration with the majority of well-known products with open API interface, and with the applications with DCOM or CORBA technologies. These systems have opportunities for connecting with customers' existing informational resources and can be transferred between different platforms and OS [4].

Disadvantages: this package isn't distributed as "box version", it is recommended to use it as custom-built product.

3.3. IBM WebSphere portal

IBM WebSphere portal is the integral part of WebSphere software platform. This platform is designed for:

– access granting for all categories of users, devices and settings options;

– integration and automation of all business processes;

– development and management of the applications.

Three classes of IBM WebSphere solutions are complied with these functional sections.
Functional framework and tools – planning, development and management of the applications. WebSphere application server, MQ package for message exchange and the most modern development tools form the powerful basis for the platform. This foundation and tools help customer to give necessary functions for Internet working, allow to develop and use of web services, provide the connectivity with scaled community of WebSphere users.
Business integration – providing of the compatibility with internal business processes. WebSphere business integrator as integral part of WebSphere package facilitate the establishment of applications and business processes including solutions for supply chain management (SCM) and integration of existing processes with web [5].
Customers' services quality – content personalization and access granting for all categories of devices. These WebSphere products are responsible for exact settings of portal in accordance with cistomers' requirements and provide wide possibilities for access of customers, business partners and ex-territorial branches.
Advantages: powerful flexible system; large repository of objects; containing of portlets for educational portals.
Disadvantages: high cost; labour-intensiveness of application development

3.4. Oracle 9i Application Server Portal

Oracle 9i Application Server has wide abilities for publishing web services and applications in Internet. One can develop web services on the basis of the modern technologies, including Java languages (with the standard J2EE), XML, PL/SQL. JDeveloper and Forms Developer as the parts of Oracle Internet Developer Suite package allow development and setup of application in the close connection with Oracle 9i Application Server.
With use of Oracle 9i AS portal and its portlet technology one can develop personified portals with one-time authorization. This action facilitates the access to the services content and provides the necessary security level. New e-business application – Webtop – delivers the personified information to different users' groups, increases their productivity and supplies with necessary data [6].

Oracle portal package can be presented as three distinct but interlinked modules: in fact, portal components, tools for development and common administration of the site and integrated tool for development of web applications based on Oracle DBMS.
Oracle portal applications consist of three different types of components:

– forms are designed for information input by use of text fields, multiply choice controls, flags, drop-down lists and dialog windows;

– reports display DBMS information in HTML, Text or Microsoft Excel formats;

– diagrams show the information in graphical mode;

– hierarchy display data with subordination status (e.g., organizational charts).

Advantages: powerful upcoming system; well-designed development tools; extensive additional abilities.
Disadvantages: too expensive system; there are few educational portlets.

3.5. Microsoft SharePoint Portal Server 2001

Microsoft SharePoint Portal Server 2001 is the flexible solution for development of portals that facilitate the search, shared working and publishing of information.
SharePoint Portal Server allows create the powerful work-ready portal systems. It consolidates main functions for document circulation management such as extraction of the necessary documents and its return to the repository, documents' profiles and shared publication. Users can create the special working area of SharePoint Portal Server that may implement project planning and document management and sharing. Electronic panels node of MS SharePoint Portal Server executes the functions of centralized access point to the information from different sources. At the same time it provides the security of used documents [6]. External data sources may be used as working areas of SharePoint Portal Server, MS Exchange Server 5.5 shared folders, local and network filesystems.
Microsoft SharePoint Portal Server special features:

– access facilitation to the business information (indexing of most important data sources; supporting of security settings for Windows and Exchange server);

– close integration with Microsoft Office products and Windows Active Desktop (abilities for extraction and returning of the documents, version control);

– effective solutions for development and enlargement of the portal (portal interface built at the basis of electronic instrumental panels allows its extension by use of webcomponents; standardized portal platform on the basis of XML, WebDAV, ADO/OLEDB and CDO standards).

Advantages: powerful complex solution oriented exclusively to the own MS-platform.
Disadvantages: there are no specified portal solutions for educational tasks; unsatisfactory parameters of security, dependability, control, support of standards.

3.6. Freeware and problem-oriented software

At analysis of different portal systems one should pay attention to the numerous solutions of freeware and shareware products, supported by groups of enthusiasts. As examples Zope [7], PHP-Nuke [8], MetaDot Portal Server [9] may be examined. These products can't be recommended as the federal-wide (nation-wide) solutions. They may be established as the school or university platform only.
There are several software products responsible for carrying out distinct useful functions: WebCT [10], Lotus LearningSpace [11]. Such kind products can be integrated to the portal as the "independent" application but not as the basis for it [12].
Advantages: free of charge; simple installation and usage; originality of solutions.
Disadvantages: absence of high-quality documentation; no guarantees; numerous errors and defects.

4. Conclusions

X-Ware platform may be recommended to the development of distributed system of educational resources' catalogue in accordance with common search engine and rubricators of educational portals system in case of brining to collaboration manufacture company.
IBM WebSphere platform is the best choice among well-known and supported products for educational portal system. It complies with complex requirements and supports the integration of existing resources into a single whole. Its usage may be combined with X-Ware implementation for distributed catalogues.
Oracle products may be recommended for realization of distinct portal modules responsible for data storing and processing.
Freeware products (Zope, PHP-Nuke, etc.) are intended for faculty-level or university-level projects.
The most important factor for the choice of the platform for portal is the presence of manufacturer's experience of working with the given platform.

References

[1] M. Bulgdakov, "Educational portal: requirements and platforms", *Elearn. World*, no. 3, pp. 38–54, 2004.

[2] A. Zhuravliov, "International technological standards of distance learning", *Acad. Dist. Edu. Bull.*, no. 1, pp. 45–49, 2003.

[3] T. Kindberg, J. Barton, and J. Morgan, "People, places, things: web presence for the real world", HPLabs Techn. Rep. HPL-2000-16, 2000.

[4] K. Dey, G. Abowd, and D. Salber, "A context-based infrastructure for smart environment", in *Proc. 1st Int. Worksh. Manag. Interact. Smart Envir.*, Dublin, Ireland, 1999, pp. 13–14.

[5] T. Kelly, "Thin-client web access patterns: measurements from cache-busting proxy", *Comput. Commun.*, vol. 25, no. 4, pp. 357–366, 2002.

[6] E. Felten and M. Schneider, "Timing attacks on web privacy", in *Proc. 7th ACM Conf. Comput. Commun. Secur.*, Athens, Greece, 2000, pp. 310–316.

[7] Zope Community, http://zope.org

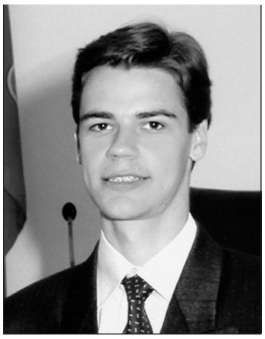
[8] PHP-Nuke Hosting Resources, http://php-nuke.org

[9] Metadot – Open Source Portal Server, http://www.metadot.com

[10] WebCT – A Blackboard Company, http://www.webct.com

[11] IBM Lotus Software, http://lotus.com

[12] "IEEE Guide for Developing User Open System Environment (OSE) Profiles", IEEE Std. 1003.23-1998(Withdrawn).



Artem Boyarchuk received his B.Sc. in 2004, M.Sc. in 2005 and then has begun his Ph.D. studies in computer engineering. Since 2003 he has worked as a specialist of Laboratory of Distance Learning, National Aerospace University "KhAI" coordinating of regional distance learning centres, development of teaching materials for distance learning and administering of informational centre of development and implementation of European Union educational projects. He also has been an adviser on International Relations of Director General of Academy of Distance Education (ADE) and workgroup coordinator on development of ADE web portal. Since 2005 he has worked as a senior specialist of Department of International projects and relations, National Law Academy of Ukraine. He has been a team leader on development of logistic concept of Library Resource Centre. His main interest focuses on issues connected with distance learning. He is an author of 6 scientific articles, 1 training supply, 1 joint UK-Ukrainian technical report on development of dependable web services.

e-mail: a.boyarchuk@khai.edu
Distance Learning Laboratory
National Aerospace University "KhAI"
Chkalov st 17
Kharkov, 61070, Ukraine

Regular paper

Transmitter diversity antenna selection techniques for wireless channels utilizing differential space-time block codes

Le Chung Tran, Tadeusz A. Wysocki, Alfred Mertins, and Jennifer Seberry

Abstract— The paper deals with transmitter diversity antenna selection techniques (ASTs) for wireless channels utilizing differential space-time block codes (DSTBCs). The proposed ASTs tend to maximize the signal-to-noise ratio (SNR) of those channels. Particularly, we propose here the so-called *general* $(M,N;K)$ AST/DSTBC scheme for such channels. Then, based on this AST, we propose two modified ASTs which are more amenable to practical implementation, namely the *restricted* $(M,N;K)$ AST/DSTBC scheme and the $(N+\bar{N},N;K)$ AST/DSTBC scheme. The *restricted* $(M,N;K)$ AST/DSTBC scheme provides relatively good bit error performance using only one feedback bit for transmission diversity purpose, while the $(N+\bar{N},N;K)$ AST/DSTBC scheme shortens the time required to process feedback information. These techniques remarkably improve bit error rate (BER) performance of wireless channels using DSTBCs with a limited number (typically 1 or 2) of training symbols per each coherent duration of the channel. Simulations show that the proposed AST/DSTBC schemes outperform the DSTBCs without antenna selection even with only 1 training symbol.

Keywords—differential space-time modulation, differential space-time block codes, diversity antenna selection, MIMO.

1. Introduction

The diversity combination of space-time codes (STCs) and a closed loop antenna selection technique (AST) assisted by a feedback channel to improve the performance of wireless channels in multiple input multiple output (MIMO) systems has been intensively examined in literature for the case of *coherent detection*, such as [5–10]. However, ASTs for channels utilizing differential space-time block codes (DSTBCs) with *differential detection* have not been considered yet. The backgrounds on DSTBCs can be found in [11–17].

In this paper[1], we propose some ASTs which tend to maximize the signal-to-noise ratio (SNR) for the channels using DSTBCs with arbitrary number M of transmit antennas (MTx antennas) and with K receive antennas (KRx antennas). Particularly, we first propose an AST called the *general* $(M,N;K)$ AST/DSTBC where the transmitter selects NTx antennas out of MTx antennas ($M > N$) to maximize the channel SNR. The antenna selection (at the transmitter) is based on the results of the comparison carried out (at the receiver) between the instantaneous powers of signals which are received during the initial transmission. The *general* $(M,N;K)$ AST/DSTBC significantly improves the performance of channels using DSTBCs. However, when M and N grow large, the number of feedback bits required to inform the transmitter also grows large. This drawback impedes the *general* $(M,N;K)$ AST/DSTBC from practical implementation if M and N are large.

The aforementioned drawback can be overcome by either reducing the number of feedback bits or shortening the time required to process feedback information. Based on these observations, we modify the *general* $(M,N;K)$ AST/DSTBC and derive the two following ASTs which are more amenable to practical implementation.

First, we propose the so-called *restricted* $(M,N;K)$ AST/DSTBC, which provides good bit error performance using only 1 feedback bit for transmission diversity purpose.

Then, we describe the so-called $(N+\bar{N},N;K)$ AST/DSTBC which shortens the average time required to process feedback information in comparison with the *general* $(M,N;K)$ AST/DSTBC, where $M = N+\bar{N}$. This AST is first motivated by the $(N+1,N;K)$ AST/STBC which we mentioned in [1] for channels using space-time block codes (STBCs) with *coherent detection*. The background on STBCs can be found in [18–21].

We show that DSTBCs associated with the proposed ASTs provide much better bit error performance than that without antenna selection. The proposed ASTs in this paper are the generalization of our ASTs published in [2, 3]. The content of this paper is also somewhat related to our published papers [1, 4].

Although, the authors propose here the ASTs for a very general case, where the system contains arbitrary numbers of Tx and Rx antennas, it is important having in mind that it is more practical to have diversity antennas installed at the transmitter, e.g., a base station in mobile communication systems, rather than at the hand-held, tiny receiver, such as a mobile phone. It is well known that the installation of more than 2Tx antennas in mobile phones is almost impractical due to the battery life-time and the small size of the phones.

Consequently, by using the term *antenna selection* in this paper, we mean *transmitter diversity* antenna selection, rather than receiver diversity antenna selection, i.e., all KRx antennas are used without selection (although the generalization of the proposed ASTs to receiver diversity antenna selection is straightforward). It should be also

[1]Related to the content in this paper are the published works [1–4].

noted that the term *differential space-time block codes* (DSTBCs) used throughout this paper means *complex, orthogonal* DSTBCs.

This paper is organized as follows.

Section 2 reviews the conventional DSTBCs mentioned in literature and provides some remarks on the time-varying Rayleigh fading channels where DSTBCs can be practically used. In Section 3, we mention some notations and assumptions used throughout this paper. Section 4 starts with the discussion on the criterion of antenna selection in channels using STBCs and then analyzes our modifications to apply to channels using DSTBCs. In Section 5, we propose the *general* $(M,N;K)$ AST/DSTBC. In Section 6.1, we propose the *restricted* $(M,N;K)$ AST/DSTBC. The $(N+\bar{N},N;K)$ AST/DSTBC is proposed in Section 6.2. Section 7 provides the mathematical expression of the relative time reduction gained by the $(N+\bar{N},N;K)$ AST/DSTBC in comparison with the *general* $(M,N;K)$ AST/DSTBC. In Section 8, we give some comments on the spatial diversity order of our proposed ASTs. Simulation results are presented in Section 9 and the paper is concluded by Section 10.

2. Reviews on DSTBCs

In this section, we review the conventional DSTBCs mentioned in literature and provide some remarks on the time-varying Rayleigh fading channels where DSTBCs can be practically used. This section is indispensable in order for the readers to understand what has been modified in the transmission procedures of DSTBCs in our proposed ASTs. It is also vital for the readers to notice the underlying requirement of all conventional DSTBCs that the channel coefficients must be constant during at least two consecutive code blocks. We also show here in which scenarios DSTBCs (differential detection) should be used instead of STBCs (coherent detection).

2.1. Conventional DSTBCs without diversity antenna selection

Differential space-time block codes are the candidate for the channels where fading changes so fast that the transmission of the training signals (eg., a large overhead) is either impractical or uneconomical. DSTBCs have been considered intensively and a number of DSTBCs have been proposed in literature such as [11–17]. In [2, 3], we have proved that all conventional DSTBCs (without antenna selection) provide a full spatial diversity order.

Let us consider the unitary DSTBC proposed by Ganesan *et al.* in [13] as an example. We consider a system with NTx antennas and KRx antennas. Let $\mathbf{R}_t$, $\mathbf{A}$, $\mathbf{N}_t$ be the $(K\times N)$-sized matrices of received signals at time t, channel coefficients between Rx and Tx antennas, and noise at the Rx antennas, respectively. The $\kappa\eta$th element of $\mathbf{A}$, namely $a_{\kappa\eta}$, is the channel coefficient of the path between the ηth Tx antenna and the κth Rx antenna. Channel coefficients are assumed to be identically independently distributed (i.i.d.) complex, zero-mean Gaussian random variables. Noises are assumed to be i.i.d. complex Gaussian random variables with the distribution $\mathcal{CN}(0,\sigma^2)$.

Let $\{s_j\}_{j=1}^{p}=\{s_j^R+is_j^I\}_{j=1}^{p}$ (where i^2=-1, s_j^R and s_j^I are the real and imaginary parts of s_j, respectively) be the set of p symbols, which are derived from a unitary power signal constellation S and transmitted in the tth block. Consequently, each symbol has a unitary energy, i.e., $|s_j|^2=1$.

We define a matrix $\mathbf{Z}_t=\frac{1}{\sqrt{p}}\sum_{j=1}^{p}(\mathbf{X}_j s_j^R+i\mathbf{Y}_j s_j^I)$, where the square, order-N weighting matrices $\{\mathbf{X}_j\}_{j=1}^{p}$ and $\{\mathbf{Y}_j\}_{j=1}^{p}$ are orthogonal themselves and they satisfy the permutation property. These weighting matrices are considered as the amicable orthogonal designs (AODs). The backgrounds on AODs can be found in [22]. The coefficient $\frac{1}{\sqrt{p}}$ is to guarantee that $\mathbf{Z}_t$ is a unitary matrix, i.e., $\mathbf{Z}_t\mathbf{Z}_t^H=\mathbf{I}$.

For illustration, the Alamouti DSTBC corresponding to $N=2$ is defined as

$$\mathbf{Z}_t=\frac{1}{\sqrt{2}}\begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} . \tag{1}$$

A DSTBC corresponding to $N=4$ is given below:

$$\mathbf{Z}_t=\frac{1}{\sqrt{3}}\begin{bmatrix} s_1 & s_2 & s_3 & 0 \\ -s_2^* & s_1^* & 0 & s_3 \\ -s_3^* & 0 & s_1^* & -s_2 \\ 0 & -s_3^* & s_2^* & s_1 \end{bmatrix} . \tag{2}$$

The transmission starts with an initial, identity, order-N matrix $\mathbf{W}_0=\mathbf{I}_N$ carrying no information. The matrix transmitted at time t $(t=1,2,3\ldots)$ is given by

$$\mathbf{W}_t=\mathbf{W}_{t-1}\mathbf{Z}_t . \tag{3}$$

As $\mathbf{Z}_t$ is a unitary matrix, the matrix $\mathbf{W}_t$ is also a unitary one. The model of the channel at time t, for $t=0,1,2\ldots$, ($t=0$ means the transmission of the first block $\mathbf{W}_0$, i.e., the initial transmission) is:

$$\mathbf{R}_t=\mathbf{A}\mathbf{W}_t+\mathbf{N}_t . \tag{4}$$

In all propositions of conventional DSTBCs, the channel coefficients must be constant during *at least* two adjacent code blocks[2], i.e., constant during *at least* $2N$ symbol time slots (STSs). It means that if the channel coefficient matrix $\mathbf{A}$ is assumed to be constant over two consecutive blocks $t-1$ and t, the maximum likelihood (ML) detector for the symbols $\{s_j\}_{j=1}^{p}$ is calculated as follows [13, 23]:

$$\{\hat{s}_j\}_{j=1}^{p}=\mathrm{Arg}\left\{\max_{\{s_j\},s_j\in S}\mathrm{Re}\{\mathrm{tr}(\mathbf{R}_t^H\mathbf{R}_{t-1}\mathbf{Z}_t)\}\right\} , \tag{5}$$

where Arg$\{.\}$ denotes the argument operation, tr$(.)$ denotes the trace operation, Re$\{.\}$ and Im$\{.\}$ denote the real and the imaginary parts of the argument, respectively.

If we denote T_c to be the average coherent time of the channel which represents the time-varying nature of the channel, then the channel is considered to be constant during

[2]This means that the channel coefficients are constant during each window of at least two consecutive code blocks and windows do not overlap each other.

this time. Therefore, after each duration T_c, the transmitter restarts the transmission and transmits a new initial block $\mathbf{W}_0$ followed by other code blocks $\mathbf{W}_t$ $(t = 1,2,3\ldots)$. These procedures are repeated until all data are transmitted.

Due to the orthogonality of DSTBCs, the transmitted symbols are decoded separately, rather than jointly. Therefore, if we denote:

$$D_j = \mathrm{Re}\{\mathrm{tr}(\mathbf{R}_t^H \mathbf{R}_{t-1}\mathbf{X}_j)\} + i\mathrm{Re}\{\mathrm{tr}(\mathbf{R}_t^H \mathbf{R}_{t-1} i\mathbf{Y}_j)\} \quad (6)$$

then the ML detector for the symbol s_j is [2, 3]:

$$\hat{s}_j = \mathrm{Arg}\{\max_{s_j \in S} \mathrm{Re}\{D_j^* s_j\}\}, \quad (7)$$

where D_j^* is the conjugate of D_j.

Expressions (6) and (7) show that the detection of the symbol s_j is carried out without the knowledge of channel coefficients. Particularly, the symbol s_j can be decoded by using the received signal blocks in the two consecutive transmission times, *provided that the channel coefficients are constant during two consecutive code blocks* (otherwise, we will not have the decoding expressions (5) and (7)).

The requirement that the channel coefficients must be constant during at least 2 consecutive code blocks can be relaxed if the linear prediction is used at the receiver. In this scenario, the receiver uses multiple previous received code blocks $\mathbf{R}_{t-1}, \mathbf{R}_{t-2}$, etc., to predict the relation between the current channel coefficient matrix, say $\mathbf{A}_t$, and the previous channel coefficient matrices. This approach has been mentioned in [24]. Certainly, the penalty of this approach is the complexity of the receiver structure.

It has been proved in our paper [2, 3] that all conventional DSTBCs (without ASTs) provide a full diversity of order NK, where N and K are the number of Tx and Rx antennas, respectively. We also can realize this observation in Section 4 of this paper.

2.2. Remarks on the time-varying Rayleigh fading channels

According to the frequency of channel coefficient changes, we distinguish three typical scenarios which are usually examined in practice and present the most common, real propagation conditions (see [17, p. 13] and [25, p. 2]).

1. Channel coefficient matrix $\mathbf{A}$ is random and its entries change randomly at the beginning of each symbol time slot (STS) and are constant during one STS. This scenario is referred to as the *fast* Rayleigh flat fading channel.
2. $\mathbf{A}$ is random and its entries change randomly after a duration containing a number of STSs. This scenario is referred to as the *block* Rayleigh flat fading channel. The example of this scenario will be mentioned later.
3. $\mathbf{A}$ is random but is selected at the beginning of transmission and its entries keep constant all the time. This scenario is referred to as the *slow* or *quasi-static* Rayleigh flat fading channel. Local area networks (LANs) or wide local area networks (WLANs) with a slow fading rate and a high data rate are the examples of the *quasi-static* Rayleigh flat fading channels, where the channel coefficients may be constant during thousands of STSs.

Given the above clarifications, we have the following important note. Owing to the condition that channel coefficients must be *constant during, at least, two consecutive code blocks*, in all conventional DSTBCs mentioned in literature, the channels are considered as block fading channels, although the coherent time of the channels in the case of DSTBCs (with differential detection) are much shorter than that in the case of STBCs (with coherent detection).

To illustrate, for the case of the Alamouti DSTBC, the channel coefficients must be constant during at least 4 STSs. During the first two STSs, the initial, order-2, identity matrix $\mathbf{I}_2$ which carries no information is transmitted. During the next two STSs, the Alamouti code carrying 2 symbols is transmitted. This note clarifies how fast fading channels may change when DSTBCs is utilized. Certainly, a longer coherent duration of the channel results in a more efficient utilization of DSTBCs.

We give 2 examples of block Rayleigh fading channels where coherent STBCs or DSTBCs can be used.

Example 1: We consider the scenario where the Alamouti STBCs with coherent detection can be used for the cellular mobile system with the carrier frequency $F_c = 900$ MHz. Speed of the mobile user is $v = 5$ km/h (walking speed) and the STS is assumed to be $T_s = 0.125$ ms (equivalently, the baud rate is $F_s = 8$ KBd/s. Denote $\mathrm{c} = 3.10^8$ m/s to be light speed. The maximum Doppler frequency is then calculated as

$$f_m = vF_c/\mathrm{c} = 4.17 \text{ Hz}.$$

The average coherent time T_c of the channel is estimated by the following empirical expression [26, p. 204]:

$$T_c = \frac{0.423}{f_m} = 101.52 \text{ ms}.$$

It means that the channel coefficients can be considered to be constant during almost $T_c/T_s \approx 812$ consecutive STSs, i.e., approximately 406 consecutive Alamouti code blocks. In this case, the channel coefficients change so slow that the training signals can be transmitted. In other words, STBCs with coherent detection are preferred than DSTBCs with differential detection.

Example 2: We consider another scenario where the Alamouti DSTBCs with different detection can be used for the cellular mobile system with the carrier frequency $F_c = 900$ MHz. Speed of the mobile user is $v = 60$ km/h (vehicular speed) and the STS is assumed to be $T_s = 0.5$ ms corresponding to the baud rate $F_s = 2$ KBd/s. The maximum Doppler frequency is then calculated as

$$f_m = vF_c/\mathrm{c} = 50 \text{ Hz}.$$

Similarly, the average coherent time T_c of the channel is estimated as [26, p. 204]

$$T_c = \frac{0.423}{f_m} = 8.46 \text{ ms} .$$

It means that the channel coefficients can be considered to be constant during $T_c/T_s \approx 16$ consecutive STSs, i.e., 8 consecutive Alamouti code blocks. The channel is a block Rayleigh fading one where DSTBCs can be employed. In this case, it is either impractical or uneconomical to use STBCs with coherent detection since the coherent time is too short to transmit multiple training symbols in order for the receiver to estimate the channel coefficients.

3. Definitions, notations and assumptions

For ease of exposition, we define some notations as follows.

Definition 1: $\mathcal{F}$ is defined as an order-N operation on M non-negative, real numbers $\{\varepsilon_1, \ldots, \varepsilon_M\}$ where the N indices ($N < M$) corresponding to the N largest values out of M values $\{\varepsilon_1, \ldots, \varepsilon_M\}$ are selected. We denote this operation as $\mathcal{F}_N(\varepsilon_1, \ldots, \varepsilon_M)$. The output of the operation $\mathcal{F}$ is the set of N indices which is denoted by $\hat{\mathfrak{I}}_N$.

Example 3: $M = 3$, $N = 2$, $\varepsilon_1 = 10$, $\varepsilon_2 = 20$ and $\varepsilon_3 = 30$. We have:

$$\hat{\mathfrak{I}}_2 = \mathcal{F}_2(\varepsilon_1, \varepsilon_2, \varepsilon_3) = \{2, 3\} .$$

The elements of the set $\hat{\mathfrak{I}}_2$ are the indices of ε_2 and ε_3, which are in turns the 2 largest values among $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$.

Definition 2: We define the $(M,N;K,L)$ AST/DSTBC scheme to be the transmitter and receiver diversity antenna selection technique for channels using DSTBCs with differential detection where NTx antennas are selected out of MTx antennas ($N < M$), while LTx antennas are selected out of KRx antennas ($L < K$) for transmission.

Given that notation, the $(M,N;K)$ AST/DSTBC scheme refers to as the *transmitter* diversity antenna selection technique for channels using DSTBCs with differential detection where NTx antennas are selected out of MTx antennas ($N < M$) for transmission. All KRx antennas are used without selection. Similarly, the $(M;K,L)$ AST/DSTBC scheme refers to as the *receiver* diversity AST for channels using DSTBCs where LTx antennas are selected out of KRx antennas for transmission, while MTx antennas are used without selection.

In the paper, we mainly focus on the transmitter diversity AST, i.e., the $(M,N;K)$ AST/DSTBC schemes. We sometimes compare the proposed $(M,N;K)$ AST/DSTBC schemes with the respective schemes in channels which use STBCs with coherent detection. Hence, similarly, we use the notation $(M,N;K)$ AST/STBC to refer to the transmitter diversity AST for channels using STBCs with coherent detection.

For example, if $M = 4$, $N = 2$ and $K = 1$, then the (4,2;1) AST/DSTBC is the AST where the 2Tx antennas are selected (depending on certain criteria) from 4Tx antennas for transmission, while the receiver has 1Rx antenna.

Some assumptions considered in the paper are given below.

Assumption 1: The channel coefficients between the transmitter and receiver antennas are assumed to be i.i.d. complex, zero-mean Gaussian random variables. Noises are assumed to be i.i.d. complex Gaussian random variables with the distribution $\mathcal{CN}(0, \sigma^2)$. These assumptions are applicable when the Tx and Rx antennas are sufficiently separated from one another (by a multiple of half of the wavelength) so that the Tx (and Rx) antennas are uncorrelated. The scenario where the antennas are correlated will be examined in our future works.

Assumption 2: Although channels with differential detection change faster than those with coherent detection, so that the transmission of multiple training signals is uneconomical (and, consequently, the utilization of DSTBCs is useful), we make a reasonable assumption that it is possible to transmit a few feedback bits (for each channel coherent duration T_c) from the receiver to the transmitter via a feedback channel *with a certain feedback error rate*. The feedback error rate is typically assumed to be 4% to 10%.

Finally, we want to stress the following important remarks.

Remark 1: Due to the tiny size of the receivers, such as the hand-held mobile phones in the cellular mobile systems, it is well known that employment of more than 2Tx antennas at the receiver is uneconomical. Hence, the receiver diversity antenna selection is not considered in this paper, although the generalization of the proposed techniques for the receiver diversity antenna selection is straightforward.

Remark 2: We use the modified notation $(N + \bar{N}, N; K)$ AST/DSTBC, rather than $(M,N;K)$ AST/DSTBC, where $M = N + \bar{N}$, to refer to our 3th proposed AST/DSTBC scheme in this paper. The main purpose of using this notation is to stress that $\bar{N}$Tx antennas among $(N + \bar{N})$ available Tx antennas are the standby Tx antennas. These standby Tx antennas are only used in certain conditions stipulated by the selection criteria. Those selection criteria will be mentioned in more details later.

4. Basis of transmitter antenna selection for channels using DSTBCs

In our papers [2, 3], we have proved that all conventional DSTBCs mentioned in literature, such as [13–16], provide a full spatial diversity order. This means that, if the channel contains NTx and KRx antennas, then square, order-N DSTBCs provide a full spatial diversity of order NK provided that the DSTBCs have a full rank.

Let us consider the unitary DSTBCs mentioned in Section 2.1 for instance. It is shown in [2, Eq. (11)], [3, Eq. (9)],

[12] and [23, Eq. (5.30)], that the SNR of the statistic D_j in Eq. (6) is approximately:

$$SNR_{diff} \approx \frac{\|\mathbf{A}\|_F^2}{2p\sigma^2} = \frac{\mathrm{tr}(\mathbf{A}^H\mathbf{A})}{2p\sigma^2} = \frac{\sum_{\eta=1}^{N}\left[\sum_{\kappa=1}^{K}|a_{\kappa\eta}|^2\right]}{2p\sigma^2}, \quad (8)$$

where $\|\mathbf{A}\|_F$ is the Frobenius norm of the matrix $\mathbf{A}$. Clearly, SNR has $2NK$ freedom degrees. As a result, the unitary DSTBC considered provides a full spatial diversity of order NK.

Let $\xi_\eta \equiv \sum_{\kappa=1}^{K}|a_{\kappa\eta}|^2$ $(\eta = 1,\ldots,N)$ be the total power of signals received by KRx antennas during each STS. We can rewrite SNR_{diff} as follows:

$$SNR_{diff} \approx \frac{\sum_{\eta=1}^{N}\xi_\eta}{2p\sigma^2}. \quad (9)$$

It is obvious that greater values of ξ_ηs result in a greater SNR_{diff}.

Let us consider a system comprising MTx antennas $(M > N)$ and KRx antennas. We now want to select the N best Tx antennas out of MTx antennas so that SNR_{diff} is maximized. From Eqs. (8) or (9), to maximize SNR_{diff}, we need to maximize $\|\mathbf{A}\|_F^2$. Equivalently, the N *first maximum values* out of M values $\{\xi_1,\xi_2,\ldots,\xi_M\} = \{\sum_{\kappa=1}^{K}|a_{\kappa 1}|^2, \sum_{\kappa=1}^{K}|a_{\kappa 2}|^2,\ldots,\sum_{\kappa=1}^{K}|a_{\kappa M}|^2\}$ must be selected. In other words, the indices of the N best Tx antennas are selected by the following antenna selection criterion:

$$\hat{\mathfrak{I}}_N = \digamma_N\Big(\xi_1,\ldots,\xi_M\Big) = \digamma_N\left(\sum_{\kappa=1}^{K}|a_{\kappa 1}|^2, \sum_{\kappa=1}^{K}|a_{\kappa 2}|^2,\ldots,\sum_{\kappa=1}^{K}|a_{\kappa M}|^2\right). \quad (10)$$

Again, note that the transmitter diversity antenna selection, rather than receiver diversity antenna selection, is examined in this paper. All KRx antennas are used without antenna selection.

The selection criterion in Eq. (10) is applicable only when the channel coefficients are perfectly known at the receiver. This scenario is realistic when the channel changes so slowly that the multiple training signals can be transmitted. This scenario is commonly examined in channels using STBCs with coherent detection. The ASTs are referred to as the $(M,N;K)$ AST/STBC schemes which have been intensively considered in literature [5–10].

As oppose to coherent detection, in channels using DSTBCs with differential detection, channel coefficients change faster so that the transmission of multiple training signals is either impractical or uneconomical, and consequently, the channel coefficients are unknown at the receiver.

Therefore, the antenna selection criterion in Eq. (10) cannot be directly applied to channels using DSTBCs with differential detection. However, we will show that this criterion can be modified to apply to channels using DSTBCs with differential detection.

Particularly, we will prove later in this paper that, at high SNRs, the statistical properties, i.e., means and variances, of the received signals $r_{0\kappa\eta}$s – the elements of the matrix $\mathbf{R}_0$ received during the initial transmission – are similar to those of the channel coefficients $a_{\kappa\eta}$s. As a result, at high SNRs, maximizing $\|\mathbf{R}_0\|_F^2$ tends to be the same as maximizing $\|\mathbf{A}\|_F^2$.

Based on this observation, we propose the modified antenna selection scheme for channels using DSTBCs. The transmitter selects Tx antennas on the basis of the comparison, which is carried out once per each channel coherent duration T_c at the receiver, between the power of the signals which are received by all KRx antennas during the initial transmission (the first block $\mathbf{W}_0$).

If we denote $\hat{\mathfrak{I}}_N$ to be the set of the N indices of the Tx antennas which should be selected, then the modified antenna selection criterion for channels using DSTBCs is:

$$\hat{\mathfrak{I}}_N = \digamma_N\Big(\chi_1,\ldots,\chi_M\Big) = \digamma_N\left(\sum_{\kappa=1}^{K}|r_{0\kappa 1}|^2, \sum_{\kappa=1}^{K}|r_{0\kappa 2}|^2,\ldots,\sum_{\kappa=1}^{K}|r_{0\kappa M}|^2\right).$$

This modified selection criterion is mentioned in more details in the so-called *general* $(M,N;K)$ AST/DSTBC scheme proposed as below.

5. The general $(M,N;K)$ AST/DSTBC for channels utilizing DSTBCs

In this section, we generalize our AST/DSTBC proposed in [2, 3] for channels using DSTBCs with arbitrary numbers of Tx and Rx antennas.

Let us consider a system containing MTx antennas and KRx antennas using the unitary, square, order-N DSTBCs $(N < M)$ proposed by Ganesan *et al.* [13, 27]. Note that the proposed ASTs are also applicable to any conventional DSTBC regardless of being unitary or not.

In the following analysis, the normal, lower case letters denote scalars, the bold, lower case letters denote vectors, while the bold upper case letters denote matrices. For simplicity, we omit the superscripts indicating the different coherent durations T_cs of the channel when a certain coherent duration is being considered. The superscripts are only used when we consider different coherent durations T_cs simultaneously.

The *general* $(M,N;K)$ AST/DSTBC is proposed as follows:

- At the beginning of transmission, the transmitter sends an initial block $\tilde{\mathbf{W}}_0 = \mathbf{I}_M$ via MTx antennas, rather than sending an initial block $\mathbf{W}_0 = \mathbf{I}_N$ via

NTx antennas like in all conventional DSTBCs. This transmission is referred to as the initial transmission.

We note the change in the size of matrices compared to Eq. (4) by using the tilde mark for matrices as below:

$$\tilde{\mathbf{W}}_0 = \mathbf{I}_M ,$$
$$\tilde{\mathbf{A}} = \left[\mathbf{a_1} \ \mathbf{a_2} \ \ldots \ \mathbf{a_M} \right] ,$$
$$\tilde{\mathbf{N}}_0 = \left[\mathbf{n_{01}} \ \mathbf{n_{02}} \ \ldots \ \mathbf{n_{0M}} \right] ,$$

where $\mathbf{a_j}$ $(j = 1 \ldots M)$ is the column vector of the channel coefficients a_{ij} $(i = 1 \ldots K)$ corresponding to the channel from the jth Tx antenna to the ith Rx antenna, i.e, $\mathbf{a_j}$=$[a_{1j}, \ldots, a_{Kj}]^T$, and $\mathbf{n_{0j}}$ is the noise affecting these channels during the initial transmission, i.e., $\mathbf{n_{0j}} = [n_{01j}, \ldots, n_{0Kj}]^T$. Here, the superscript T denotes the transposition operation.

- The receiver determines the matrix $\tilde{\mathbf{R}}_0$ of received signals during the initial transmission as given below:

$$\begin{aligned} \tilde{\mathbf{R}}_0 &= \tilde{\mathbf{A}}\tilde{\mathbf{W}}_0 + \tilde{\mathbf{N}}_0 \\ &= \tilde{\mathbf{A}}\mathbf{I}_M + \tilde{\mathbf{N}}_0 \\ &= \left[\mathbf{r_{01}} \ \mathbf{r_{02}} \ \ldots \ \mathbf{r_{0M}} \right] \qquad (11) \\ &= \left[\mathbf{a_1} + \mathbf{n_{01}} \ \ \mathbf{a_2} + \mathbf{n_{02}} \ \ldots \ \mathbf{a_M} + \mathbf{n_{0M}} \right] , \end{aligned}$$

where

$$\begin{aligned} \mathbf{r_{0j}} &= \mathbf{a_j} + \mathbf{n_{0j}} \\ &= [a_{1j} + n_{01j}, \ldots, a_{Kj} + n_{0Kj}]^T \quad j = 1 \ldots M . \end{aligned}$$

- From the initial received matrix $\tilde{\mathbf{R}}_0$, the receiver determines semiblindly the N best channels based on the initial, received matrix $\tilde{R}_0$ by comparing M terms χ_j=$\|\mathbf{r_{0j}}\|_F^2$, for $j = 1 \ldots M$, i.e., comparing the total power of the signals received by all KRx antennas from the jth Tx antenna during the jth STS:

$$\chi_j = \sum_{i=1}^{K} |r_{0ij}|^2 = \sum_{i=1}^{K} |a_{ij} + n_{0ij}|^2 \qquad (12)$$

to search for the first N maximum values. In other words, the antenna selection criterion is:

$$\begin{aligned} \hat{\mathfrak{I}}_N &= \digamma_N(\chi_1, \ldots, \chi_M) \\ &= \digamma_N\left(\sum_{i=1}^{K} |r_{0i1}|^2, \sum_{i=1}^{K} |r_{0i2}|^2, \ldots, \sum_{i=1}^{K} |r_{0iM}|^2 \right) \\ &= \digamma_N\left(\sum_{i=1}^{K} |a_{i1} + n_{0i1}|^2, \sum_{i=1}^{K} |a_{i2} + n_{0i2}|^2, \ldots, \right. \\ &\qquad \left. \sum_{i=1}^{K} |a_{iM} + n_{0iM}|^2 \right) , \qquad (13) \end{aligned}$$

where $\hat{\mathfrak{I}}_N$ denotes the set of N indices of the Tx antennas which should be selected.

Without loss of generality, we assume here that these maximum values are corresponding to the first N elements in the matrix $\tilde{\mathbf{R}}_0$, i.e.,

$$\hat{\mathfrak{I}}_N = \{1, 2, \ldots, N\} .$$

Then, the receiver carries out the two following tasks:

1. The receiver informs the transmitter via a feedback channel to select the first NTx antennas to transmit data.
2. The receiver generates the matrix $\mathbf{R}_0$, which is used to decode the next code blocks, by taking the first N elements of the matrix $\tilde{\mathbf{R}}_0$, corresponding to the first N maximum values, i.e., $\mathbf{R}_0$=$\left[\mathbf{a_1} + \mathbf{n_{01}} \ \ \mathbf{a_2} + \mathbf{n_{02}} \ \ldots \ \mathbf{a_N} + \mathbf{n_{0N}} \right]$.

- The transmitter selects the NTx antennas indicated by the feedback information. In this case, the first NTx antennas are selected to transmit data. The transmission is now exactly the same as that in the system using the N first Tx antennas only.

If T_c is the average coherent time of the channel, then after each duration T_c, the transmitter restarts the transmission and transmits a new initial block $\tilde{\mathbf{W}}_0$ followed by other code blocks $\mathbf{W}_t$ $(t = 1, 2, 3 \ldots)$. The above procedures are repeated until all data are transmitted.
The transmission procedure is shown in Fig. 1. The superscripts are used to indicate the different coherent durations T_cs of the channel. The code blocks $\tilde{\mathbf{W}}_0$ are transmitted via MTx antennas in M STSs and the following blocks via NTx antennas in N STSs.

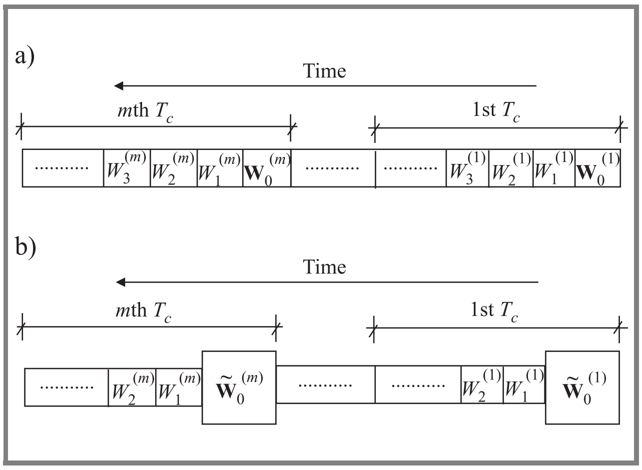


Fig. 1. Transmission of DSTBCs without (a) and with (b) the antenna selection technique.

From the aforementioned algorithm, we have following remarks.

Remark 3: At the transmitter, after the initial matrix $\tilde{\mathbf{W}}_0$ is transmitted, the next matrices $\mathbf{W}_t$ $(t = 1, 2, 3, \ldots)$ can be calculated by using a tacit default matrix $\mathbf{W}_0 = \mathbf{I}_N$ in Eq. (3). We use the term *tacit default matrix* to refer to the fact that the matrix $\mathbf{W}_0 = \mathbf{I}_N$ is tacitly used at the transmitter to generate the next code blocks $\mathbf{W}_t$ by Eq. (3), rather than being actually transmitted. Owing to this fact, it is also important to note that the generation of the matrices $\mathbf{W}_t$ does not necessarily take place after the transmitter obtains the feedback information. Instead, the next code blocks $\mathbf{W}_t$ are automatically generated by multiplying the previous block $\mathbf{W}_{t-1}$ with the tacit default matrix $\mathbf{W}_0 = \mathbf{I}_N$ following Eq. (3).

Remark 4: The above proposed AST is carried out with only $\mathcal{N}_{training} = (M - N)$ training symbols for each coherent duration T_c. The typical values of $\mathcal{N}_{training}$ are 1 or 2 symbols.

Remark 5: The number of feedback bits required to inform the transmitter about the best channels in the *general* $(M,N;K)$ AST/DSTBC is:

$$\mathcal{N} = \left\lceil \log_2 \binom{M}{N} \right\rceil, \tag{14}$$

where $\lceil . \rceil$ is the ceiling function.

Remark 6: In all conventional DSTBCs, the initial matrix $\mathbf{W}_0 = \mathbf{I}_N$ is only used to initialize the transmission. Particularly, $\mathbf{W}_0$ is used to calculate the next transmitted matrices following Eq. (3), and to generate the initial, received matrix $\mathbf{R}_0$ *directly*, which is combined with the next receiving matrix $\mathbf{R}_1$ to decode transmitted symbols.

Unlike the conventional DSTBCs without ASTs, in the proposed technique, the initial identity matrix $\tilde{\mathbf{W}}_0 = \mathbf{I}_M$ is transmitted. This matrix has two main roles. It enables the receiver to generate the initial, received matrix $\mathbf{R}_0$ *indirectly* (from the received matrix $\tilde{\mathbf{R}}_0$). Simultaneously, in some sense, it also plays a role of training signals, which assist the receiver to determine semiblindly the best channels. This is the main difference between the differential space-time coding with our AST and the one without AST.

Remark 7: Similarly to the conventional DSTBCs without ASTs mentioned in Section 3, in our proposed technique, channel coefficients are required to be constant during at least two consecutive code blocks. Therefore, the channels must be constant during, at least, $(M+N)$ STSs in our proposed AST, while they must be unchanged during at least $2N$ STSs in all conventional DSTBC techniques without the proposed ASTs if the delay of transmitting feedback information from the receiver to the transmitter is not considered. In the case when the delay is considered, the channel coefficients must stay longer.

Remark 8: The procedures of the proposed *general* $(M,N;K)$ AST/DSTBC is more explicitly presented in Fig. 2. Steps 1a, 1b, 4 and 5 are carried out at the transmitter, while the remaining steps are carried out at the receiver. As stated earlier, Step 1b is not necessarily carried out after Step 3a finishes. In other words, the transmitter can perform Step 1b right after finishing Step 1a. Similarly, because the matrix $\mathbf{R}_0$ is created straightforwardly from the matrix $\tilde{\mathbf{R}}_0$, the receiver can perform Step 3b right after finishing Step 3a. These properties reduce unnecessary delays during transmission.

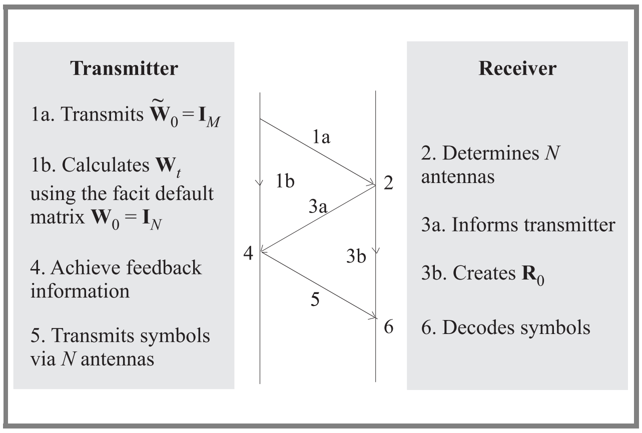


Fig. 2. The *general* $(M,N;K)$ AST/DSTBC for the system using DSTBCs.

6. The restricted $(M,N;K)$ AST/DSTBC and the $(N+\bar{N},N;K)$ AST/DSTBC

As mentioned in Eq. (14), the number of feedback bits required in the *general* $(M,N;K)$ AST/DSTBC is:

$$\mathcal{N} = \left\lceil \log_2 \binom{M}{N} \right\rceil .$$

It is easy to realize that, $\mathcal{N}$ is large for large values of M and N. For instance, in the *general* (6,4;K) AST/DSTBC (K is arbitrary), we have $\mathcal{N} = 4$. Therefore, it is either impractical or uneconomical to employ the *general* $(M,N;K)$ AST/DSTBC for large values of M and N, except when either the number of feedback bits or the time required to process feedback information is reduced.

Motivated by this observation, we derive here the two AST/DSTBC schemes which are the modifications of the aforementioned, *general* $(M,N;K)$ AST/DSTBC scheme. We refer those ASTs to as the *restricted* $(M,N;K)$ AST/DSTBC and the $(N+\bar{N},N;K)$ AST/DSTBC. The two modified ASTs are more amenable to practical implementation in channels using DSTBCs than the *general* $(M,N;K)$ AST/DSTBC.

The *restricted* $(M,N;K)$ AST/DSTBC requires only 1 feedback bit, while providing a relatively good bit error performance. Meanwhile, the $(N+\bar{N},N;K)$ AST/DSTBC requires at most an equal number of feedback bits as the *general* $(M,N;K)$ AST/DSTBC where $M = N+\bar{N}$, while shortening the time required to process feedback information. Especially, when $\bar{N} = 1$, the $(N+1,N;K)$ AST/DSTBC scheme provides the *same* bit error performance as the *general* $(M,N;K)$ AST/DSTBC scheme, where $M = N+1$, while shortening the processing time for feedback information. For $\bar{N} > 1$, there exists a degradation of the bit error performance of the $(N+\bar{N},N;K)$ AST/DSTBC scheme, compared to the *general* $(M,N;K)$ AST/DSTBC scheme where $M=N+\bar{N}$. Therefore, the $(N+1,N;K)$ AST/DSTBC scheme is of our particular interest in this paper.

6.1. The restricted (M,N;K) AST/DSTBC

In the scenario where the capacity limitation of the feedback channel, especially in the uplink channels of the 3G mobile communication systems, needs to be considered, the number of feedback bits is as small as possible. More importantly, limiting the number of feedback bits is necessary when fading changes fast. Based on the *general* $(M,N;K)$ AST/DSTBC mentioned in Section 5, we propose here the *restricted* $(M,N;K)$ AST/DSTBC for channels using DSTBCs, where only 1 feedback bit is required for each channel coherent duration T_c to inform the transmitter.

In the *restricted* $(M,N;K)$ AST/DSTBC, the set of MTx antennas is divided into two subsets. Each subset includes NTx ($N<M$) antennas. Subsets may partially overlap each other. Figure 3 presents 3 cases for illustration. In Fig. 3a,

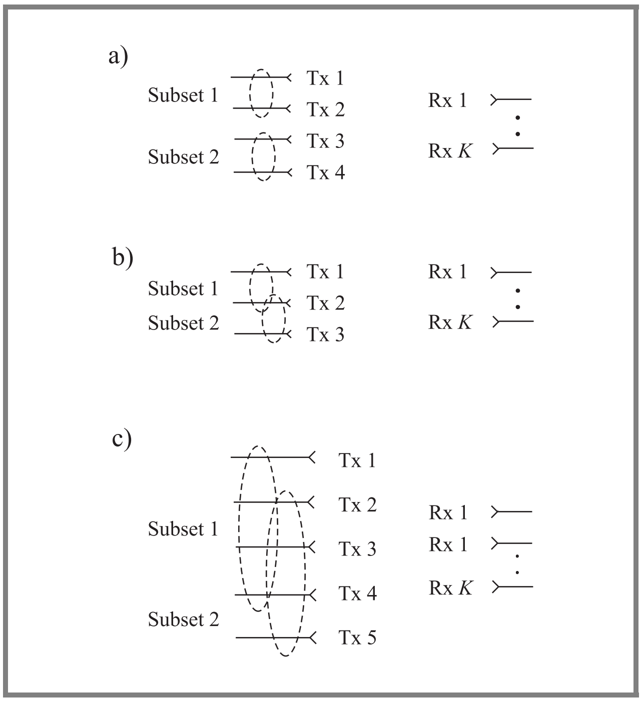


Fig. 3. Some examples of the transmitter antenna grouping for (a) the *restricted* (4,2;K) AST/DSTBC; (b) the *restricted* (3,2;K) AST/DSTBC; (c) the *restricted* (5,4;K) AST/DSTBC.

we give an example where 4Tx antennas are divided into 2 subsets including 2Tx antennas each, while in Fig. 3b, 3Tx antennas are divided into 2 subsets containing 2Tx each. These 2 cases can be applied, for instance, to the Alamouti DSTBC with the *restricted* (4,2;K) AST/DSTBC and with the *restricted* (3,2;K) AST/DSTBC, respectively. Figure 3c, we derive other example where 5Tx antennas are divided into 2 subsets which partially overlap one another and include 4Tx antennas each. This case can be applied, for instance, to the order-4 DSTBC with the *restricted* (5,4;K) AST/DSTBC.

Let Ψ and Φ be the sets of indices indicating the order of the Tx antennas in each subset, respectively. The selection criterion for the restricted $(M,N;K)$ AST/DSTBC is as follows.

During each coherent duration T_c of the channel, the receiver compares:

$$\sum_{j\in\Psi}\chi_j=\sum_{j\in\Psi}\left[\sum_{i=1}^{K}|r_{0ij}|^2\right]$$

and

$$\sum_{j\in\Phi}\chi_j=\sum_{j\in\Phi}\left[\sum_{i=1}^{K}|r_{0ij}|^2\right],$$

ie., the receiver compares the total power of the signals received by all KRx antennas during the initial transmission from two subsets of Tx antennas, and then informs the transmitter to select the subset providing the greater total power. If $\sum_{j\in\Psi}\chi_j$ is lager, then the receiver, via a feedback loop, informs the transmitter to select the Tx antennas corresponding to the set of indices Ψ. Otherwise, the Tx antennas corresponding to the set of indices Φ should be selected. These procedures are repeated for different coherent durations T_cs of the channel until the transmission of data is completed.

It is obvious that only one feedback bit per each coherent time T_c is required for transmission diversity purpose.

6.2. The $(N+\bar{N},N;K)$ AST/DSTBC

In this section, we consider a system containing $M=(N+\bar{N})$Tx antennas and KRx antennas and transmitting square, order-N DSTBCs. Among MTx antennas, NTx antennas are called default Tx antennas which are normally used to transmit signals, and $\bar{N}$ remaining Tx antennas are the standby ones which are only used when the selection criterion is satisfactory. The diagram of the system in this technique is shown in Fig. 4.

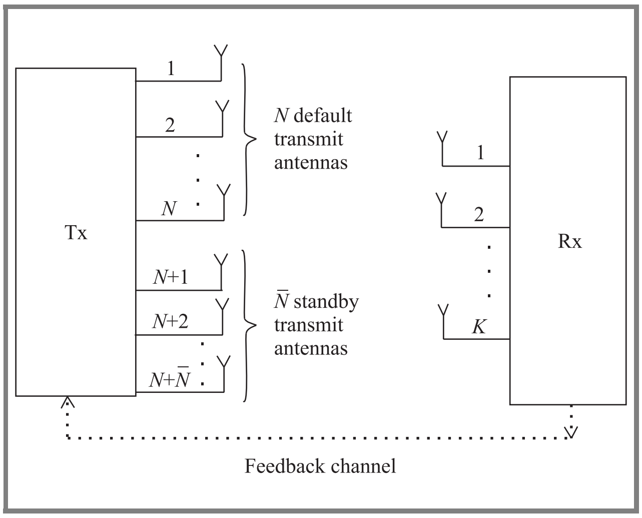


Fig. 4. The diagram of the $(N+\bar{N},N;K)$ AST/DSTBC.

We propose here a modified AST/DSTBC scheme for this structure of the system which is referred to as the $(N+\bar{N},N;K)$ AST/DSTBC. This AST shortens the time required to process feedback information in comparison with the *general* (M,N,K) AST/DSTBC where $M=N+\bar{N}$.

Note that $\bar{N}$ is *strictly smaller* than N, i.e., $\bar{N}<N$. It will be shown later that when $\bar{N}=N$, the $(N+\bar{N},N;K)$ AST/DSTBC turns into the *restricted* $(M,N;K)$ AST/DSTBC where $M=N+\bar{N}$.

Without loss of generality, we number $(N+\bar{N})$Tx antennas by indices from 1 to $(N+\bar{N})$, and assume that the N default Tx antennas are indexed from 1 to N while the $\bar{N}$ standby Tx antennas are indexed from $(N+1)$ to $(N+\bar{N})$.

Similarly to the *general* $(M,\ N,\ K)$ AST/DSTBC, in the $(N+\bar{N},N;K)$ AST/DSTBC, the transmitter starts the transmission by transmitting an identity, order-M matrices $\tilde{\mathbf{W}}_0=\mathbf{I}_M=\mathbf{I}_{N+\bar{N}}$ during each channel coherent time T_c. Let $\tilde{\mathbf{R}}_0$ be the initial, received matrix $\tilde{\mathbf{R}}_0$ during the initial transmission, i.e., the time when the initial matrix $\tilde{\mathbf{W}}_0$ is transmitted. Similarly to Eq. (11), we have:

$$\tilde{\mathbf{R}}_0=\left[\,\mathbf{r_{01}}\ \ \mathbf{r_{02}}\ \ \ldots\ \ \mathbf{r_{0N}}\ \ \mathbf{r_{0N+1}}\ \ \ldots\ \ \mathbf{r_{0N+\bar{N}}}\,\right].$$

In this expression, $\mathbf{r_{0j}}$ is the column vector of the signals received by all KRx antennas during the jth STS from the jth Tx antenna. Let χ_j=$\|\mathbf{r_{0j}}\|_F^2$ which is the total power received by all KRx antennas from the jth Tx antenna $(j=1,\ldots,N+\bar{N})$.

We denote φ_k to be the set of $\bar{N}$ indices of the $\bar{N}$ default Tx antennas which are arbitrarily taken from N default Tx antennas. There are total $q=\binom{N}{\bar{N}}$ such sets. Furthermore, for $k=1,\ldots,q$, we denote:

$$\begin{aligned}\alpha_k &= \sum_{j\in\varphi_k}\chi_j\\ &= \sum_{j\in\varphi_k}\|\mathbf{r_{0j}}\|_F^2\\ &= \sum_{j\in\varphi_k}\left[\sum_{i=1}^{K}|r_{0ij}|^2\right].\end{aligned}$$

The proposed $(N+\bar{N},N;K)$ AST/DSTBC is as follows. On the one hand, the receiver searches for the minimum value among q values $\{\alpha_1,\ldots,\alpha_q\}$. Let α be this minimum value and $\hat{\mathfrak{I}}_{\bar{N}}$ be the set of indices of the corresponding default Tx antennas. This action can be mathematically presented by

$$\alpha=\min\ \left\{\alpha_1,\ldots,\alpha_q\right\}.$$

On the other hand, the receiver calculates the total power of the received signals value which are received by all KRx antennas during the initial transmission from $\bar{N}$ standby Tx antennas. If we denote this total power to be β, then this action can be expressed as:

$$\begin{aligned}\beta &= \sum_{j=(N+1)}^{(N+\bar{N})}\chi_j\\ &= \sum_{j=(N+1)}^{(N+\bar{N})}\|\mathbf{r_{0j}}\|_F^2\\ &= \sum_{j=(N+1)}^{(N+\bar{N})}\left[\sum_{i=1}^{K}|r_{0ij}|^2\right].\end{aligned}$$

If $\alpha\geq\beta$, then the Tx antennas which the transmitter should select are all default Tx antennas $\{1,\ldots,N\}$.

If $\alpha<\beta$, the $\bar{N}$ default Tx antennas which have the indices listed in the set $\hat{\mathfrak{I}}_{\bar{N}}$ will be replaced by the standby antennas. To illustrate, we assume that $\hat{\mathfrak{I}}_{\bar{N}}=\{1,2,\ldots,\bar{N}\}$, i.e., the first $\bar{N}$ default Tx antennas provide the minimum value α. If $\alpha\geq\beta$, then the Tx antennas are $\{1,2,\ldots,N\}$. Otherwise, the first $\bar{N}$ default Tx antennas are replaced by the $\bar{N}$ standby Tx antennas. Consequently, the NTx antennas which should be selected are $\{\bar{N}+1,\ldots,N-1,N,N+1,\ldots,N+\bar{N}\}$.

The antenna selection mechanism for this example is presented more clearly by the flowchart in Fig. 5.

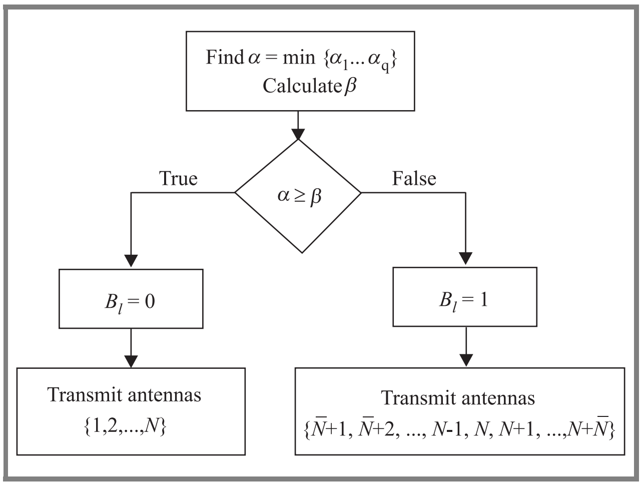


Fig. 5. The flow chart of the $(N+\bar{N},N;K)$ AST/DSTBC .

Associated with this antenna selection mechanism, we propose the structure of the feedback information as presented in Fig. 6. The bit B_l is used to indicate whether the trans-

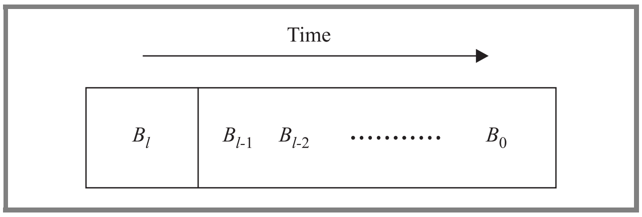


Fig. 6. The proposed structure of the feedback information for channels using DSTBCs.

mitter has to replace $\bar{N}$ default antennas with the standby ones. The bit B_l is zero if the answer is no, i.e., $\alpha \geq \beta$, and B_l is unity otherwise. The l following bits indicate which $\bar{N}$ antennas among N default antennas should be replaced by the standby ones.

It is easy to realize that $l = \lceil \log_2 \binom{N}{\bar{N}} \rceil$. With this structure, the transmitter first considers the bit B_l. As soon as it realizes that $B_l = 0$, the rest of the feedback information is not necessarily processed. The transmitter will transmit signals via the default Tx antennas $\{1, 2, \ldots, N\}$. If $B_l = 1$, the transmitter uses the l following bits $B_{l-1}, \ldots, B_0$ to recognize which default antennas should be replaced by the standby ones.

Therefore, the number of feedback bits required to be transmitted in the $(N+\bar{N},N;K)$ AST/DSTBC is at most equal to:

$$\mathcal{N}_1 = l + 1 = 1 + \left\lceil \log_2 \binom{N}{\bar{N}} \right\rceil . \qquad (15)$$

We want to stress that, theoretically, there is no need to transmit l bits $B_{l-1}, \ldots, B_0$ in the case $B_l = 0$. If so, a single feedback bit (bit B_l) is required to be transmitted.

Note that the number of feedback bits required to be transmitted and processed in the *general* $(M,N;K)$ AST/DSTBC where $M = N+\bar{N}$ is always:

$$\mathcal{N}_2 = \left\lceil \log_2 \binom{M}{N} \right\rceil = \left\lceil \log_2 \binom{N+\bar{N}}{N} \right\rceil .$$

It is easy to realize that if N is a power of 2, we have $\mathcal{N}_1 \leq \mathcal{N}_2$. For instance, for $N = 2$ and $\bar{N} = 1$, we have $\mathcal{N}_1 = \mathcal{N}_2 = 2$. For $N = 4$ and $\bar{N} = 2$, we have $\mathcal{N}_1 = 3$ and $\mathcal{N}_2 = 4$.

Therefore, if N is the power of 2, the number of feedback bits required to be transmitted in the $(N+\bar{N},N;K)$ AST/DSTBC is almost equal to that required in the *general* $(M,N;K)$ AST/DSTBC ($M = N+\bar{N}$). The number of feedback bits required to be processed in the $(N+\bar{N},N;K)$ AST/DSTBC is either $(l+1)$, which is equal to the number of transmitted feedback bits $\mathcal{N}_1$, or only 1 (smaller than $\mathcal{N}_1$) depending on the bit B_l. The smaller number of feedback bits required to be transmitted and to be processed in the $(N+\bar{N},N;K)$ AST/DSTBC shortens the time required to process feedback information in the $(N+\bar{N},N;K)$ AST/DSTBC in comparison with the time required in the *general* $(M,N;K)$ AST/DSTBC. The quantitative estimation of this time reduction will be mentioned later.

From the aforementioned algorithm, we have the following remarks on the $(N+\bar{N},N;K)$ AST/DSTBC.

Remark 9: Theoretically, it is not necessary to transmit l bits $B_{l-1}, \ldots, B_0$ in the case $B_l = 0$. Only one feedback bit B_l is required to be transmitted (and processed) in this case. This observation may further shortens the time for feeding information back.

Remark 10: The N default Tx antennas are always used for transmission whenever $\beta \leq \alpha$, i.e., the set of $\bar{N}$ standby antennas is not better[3] than the worst set of $\bar{N}$ default Tx antennas among N default Tx antennas.

When $\beta > \alpha$, i.e, the set of $\bar{N}$ standby antennas is better than the worst set of $\bar{N}$ default Tx antennas among N default Tx antennas, these $\bar{N}$ standby antennas are used to replace the $\bar{N}$ default antennas.

Remark 11: If $\bar{N} = N$, not only the antenna selection criterion of the $(N+\bar{N},N;K)$ AST/DSTBC is exactly the same as that of the *restricted* $(M,N;K)$ AST/DSTBC where $M = N+\bar{N}$, but the required numbers of feedback bits of both ASTs are also the same (only 1 feedback bit is required). Therefore, the $(N+\bar{N},N;K)$ AST/DSTBC turns into the *restricted* $(M,N;K)$ AST/DSTBC. Owing to this reason, $\bar{N}$ must be strictly smaller than N in the $(N+\bar{N},N;K)$ AST/DSTBC.

Remark 12: If $2 \leq \bar{N} < N$, the $(N+\bar{N},N;K)$ AST/DSTBC is suboptimal as the set containing the N best Tx antennas among $(N+\bar{N})$Tx antennas is not always selected for transmission, and consequently, it provides a worse BER performance than the *general* $(M,N;K)$ AST/DSTBC where $M = N+\bar{N}$. In return for this disadvantage, the $(N+\bar{N},N;K)$ AST/DSTBC shortens the time required to process feedback information in comparison with the *general* $(M,N;K)$ AST/DSTBC.

Remark 13: If $\bar{N} = 1$, the antenna selection criterion of the $(N+\bar{N},N;K)$ AST/DSTBC turns into the selection criterion of the *general* $(M,N;K)$ AST/DSTBC where $M = N+\bar{N} = N+1$. Intuitively, both the $(N+1,N;K)$ AST/DSTBC and the *general* $(M,N;K)$ AST/DSTBC select the N optimal Tx antennas out of $(N+1)$Tx antennas. Consequently, the BER performance of the $(N+1, N; K)$ AST/DSTBC is the same as that of the *general* $(M,N;K)$ AST/DSTBC $(M = N+1)$.

The main advantage of the $(N+1, N; K)$ AST/DSTBC over the *general* $(M,N;K)$ AST/DSTBC is that the time required to process feedback information in the former is shorter than that in the later. This advantage will be mentioned in more details in the next section in which the quantitative estimation of the time reduction gained by the $(N+1, N; K)$ AST/DSTBC in comparison with the *general* $(M,N;K)$ AST/DSTBC is derived.

Owing to these reasons, the $(N+\bar{N},N;K)$ AST/DSTBC with $\bar{N} = 1$, i.e., the $(N+1, N; K)$ AST/DSTBC, is of our particular interest in this paper.

Let $\chi_j = \sum_{i=1}^{K} |r_{0ij}|^2$ for $j = 1, \ldots, (N+1)$. The $(N+1, N; K)$ AST/DSTBC scheme can be slightly modified from the $(N+\bar{N},N;K)$ AST/DSTBC and stated as follows.

The receiver searches for the minimum value $\chi_{\min}$ among $(N+1)$ values $\{\chi_1, \ldots, \chi_{N+1}\}$, i.e.:

$$\chi_{\min} = \min\{\chi_1, \ldots, \chi_{N+1}\} .$$

We assume that $\chi_{\min} \equiv \chi_n$ where $n = 1, \ldots, (N+1)$.

[3]A better set provides a larger total power which is received by all KRx antennas during the initial transmission.

If $n \equiv (N+1)$, then all N default Tx antennas are used to transmit signals. In this case, bit $B_l = 0$. Otherwise, the indexed-n default Tx antenna is replaced by the standby Tx antenna (the $(N+1)$th Tx antenna). This standby antenna is combined with the $(N-1)$Tx antennas to transmit signals. In this case, bit $B_l = 1$.

7. Relative reduction of the average processing time of the $(N+\bar{N},N;K)$ AST/DSTBC

In order to estimate the time reduction obtained by the $(N+\bar{N},N;K)$ AST/DSTBC, we compare the average time required to process feedback information in this AST and that required in the *general* $(M,N;K)$ AST/DSTBC ($M = N+\bar{N}$) in Section 5.

Although, there is a fact that the time required to process the feedback information does not necessarily increase linearly with the number of feedback bits, it is easier to calculate the time benefit of the proposed technique when the average processing time is assumed to increase linearly with the number of feedback bits. Obviously, the result we derive as follows is only aimed at providing the readers with the lower bound of the relative reduction of the average processing time obtained by the $(N+\bar{N},N;K)$ AST/DSTBC in comparison with that of the *general* $(M,N;K)$ AST/DSTBC.

Let P_0 be the probability of the event that the set of $\bar{N}$ standby Tx antennas is not used in the $(N+\bar{N},N;K)$ AST/DSTBC. In other words, P_0 is the probability of the event that $\beta \le \alpha$, i.e. $P_0 = P(\beta \le \alpha)$. Similarly, let P_1 be the probability of the event that the $\bar{N}$ standby Tx antennas are used for transmission, i.e, $P_1 = P(\beta > \alpha)$. Clearly, we have $P_1 = (1-P_0)$.

We now calculate P_0 in the two following cases which are different in the underlying essences.

- When $\bar{N} = 1$, as mentioned earlier in Remark 13, the default Tx antenna is only used when it is the worst Tx antenna among $(N+1)$Tx antennas. We make a reasonable assumption that the event where a certain Tx antenna (either default or standby antenna) is the worst antenna among $(N+1)$Tx antennas is equiprobable. Then we have:

$$P_0 = P(\beta \le \alpha) = \frac{1}{\binom{N+1}{1}} = \frac{1}{(N+1)} . \quad (16)$$

- When $\bar{N} \ge 2$, we make a reasonable assumption that the event in which a set containing the certain $\bar{N}$ default Tx antennas selected from the N available default Tx antennas is the worst set, is equiprobable. This means that:

$$P(\alpha \equiv \alpha_1) = \cdots = P(\alpha \equiv \alpha_q) = \frac{1}{\binom{N}{\bar{N}}} = \frac{1}{q} .$$

We also assume the following conditional probability:

$$P(\beta \le \alpha | \alpha \equiv \alpha_k) = 0.5$$

for $k = 1,\ldots,q$. As a result, we have:

$$\begin{aligned} P_0 &= P(\beta \le \alpha) \\ &= \sum_{k=1}^{q} P(\beta \le \alpha | \alpha \equiv \alpha_k) P(\alpha \equiv \alpha_k) \\ &= \sum_{k=1}^{q} 0.5 \cdot \frac{1}{q} \\ &= 0.5 . \end{aligned} \quad (17)$$

Let ϑ be the average processing time for 1 feedback bit. Because the transmitter has to process 1 feedback bit (bit B_l) only if $\beta \le \alpha$ and has to process $\mathcal{N}_1 = \left(1 + \left\lceil \log_2 \binom{N}{\bar{N}} \right\rceil\right)$ feedback bits if $\beta > \alpha$, the average time required to process feedback information in the $(N+\bar{N},N;K)$ AST/DSTBC is:

$$\begin{aligned} \tau_1 &= P_0 \vartheta + P_1 \mathcal{N}_1 \vartheta \\ &= P_0 \vartheta + (1-P_0)\left(1 + \left\lceil \log_2 \binom{N}{\bar{N}} \right\rceil\right) \vartheta . \end{aligned}$$

On the other hand, in the *general* $(M,N;K)$ AST/DSTBC where $M = N+\bar{N}$, the transmitter always has to process $\mathcal{N}_2 = \left\lceil \log_2 \binom{N+\bar{N}}{N} \right\rceil$ feedback bits. Therefore, the average processing time is:

$$\tau_2 = \mathcal{N}_2 \vartheta = \left\lceil \log_2 \binom{N+\bar{N}}{N} \right\rceil \vartheta .$$

Hence, the relative reduction of the average processing time between two techniques is:

$$\begin{aligned} \frac{\triangle\tau}{\tau_2} &\triangleq \frac{\tau_2 - \tau_1}{\tau_2} \\ &= 1 - \frac{1 + (1-P_0)\left\lceil \log_2 \binom{N}{\bar{N}} \right\rceil}{\left\lceil \log_2 \binom{N+\bar{N}}{\bar{N}} \right\rceil} . \end{aligned} \quad (18)$$

For $\bar{N} = 1$, from Eqs. (16) and (18), we have:

$$\frac{\triangle\tau}{\tau_2} = 1 - \frac{1 + (1 - \frac{1}{N+1})\left\lceil \log_2 N \right\rceil}{\left\lceil \log_2 (N+1) \right\rceil} .$$

For $\bar{N} \ge 2$, from Eqs. (17) and (18), we have:

$$\frac{\triangle\tau}{\tau_2} = 1 - \frac{1 + 0.5\left\lceil \log_2 \binom{N}{\bar{N}} \right\rceil}{\left\lceil \log_2 \binom{N+\bar{N}}{\bar{N}} \right\rceil} .$$

The relative time reduction $\frac{\triangle\tau}{\tau_2}$ [%] for some particular values of N and $\bar{N}$ is presented by the table in Fig. 7. We only need to calculate the time reduction for the pair of N and $\bar{N}$ satisfying $\bar{N} < N$.

From this table, we realize that the average processing time reduction is considerable even for $\bar{N} = 1$. In this case, the average processing time reduction for $N = 2$, 4 and 8 is 16.67, 13.33 and 8.33%, respectively. To illustrate, the (2+1, 2; 1) AST/DSTBC in the system using the Alamouti DSTBC with 2 default Tx antennas, 1 standby Tx antenna and 1Rx antenna gains the relative time reduction of 16.67%.

$\bar{N}$ \ N	2	4	8
1	16.67	13.33	8.33
2		37.5	41.67
3		66.67	50
4			50

Fig. 7. Relative time reduction [%] of the $(N+\bar{N},N;K)$ AST/DSTBC compared to the *general* $(M,N;K)$ AST/DSTBC where $M=N+\bar{N}$.

It is worth to stress that the time reduction is probably much greater than the above figures if we take its non-linear proportionality with the number of feedback bits into consideration.

8. Some comments on spatial diversity order of the proposed ASTs

In this section, we consider the spatial diversity order of the ASTs proposed for channels using DSTBCs with differential detection. To do that, at first, we review the same issue for channels using STBCs with coherent detection, to provide the readers with the state of the art of this issue.

The spatial diversity order of the ASTs for channels using space-time codes with coherent detection has been somewhat examined in a few papers, such as [8–10, 28–31]. Particularly, in [28] and [30], the authors considered the combination of the transmitter antenna selection and space-time trellis codes (STTCs) and proved that the (M,2;1) AST/STTC and (M,2;2) AST/STTC schemes provide a full spatial diversity order when SNR is very large (see Eqs. (26) and (27) in [28]) as long as the STTCs have a full rank. In [31], the authors considered the receiver (not transmitter) diversity selection associated with the use of STCs (either STBCs or STTCs) in MIMO systems over the quasi-static (slow) Rayleigh fading channels. The author proved there that the $(M;K,L)$ AST/STC schemes (where MTx antennas are used without selection, while the L best Tx antennas are selected out of KRx antennas) provide a full spatial diversity order of MK, provided that the STCs have a full rank (see Eq. (10) in [31]).

It is noted that, in this paper, we consider the *transmitter* (*not receiver*) diversity antenna selection and the use of DSTBCs which have *orthogonal* structures. Therefore, it is useful to review the spatial diversity order of the ASTs associated with STBCs only (not STTCs or other STCs). Having this note in mind, we realize that there are very few works, such as [10] and [29], have mentioned the spatial diversity order of transmitter diversity ASTs for channels using STBCs. In [10] and [29], the authors limited themselves to consider the Alamouti STBC modulated by a binary phase shift keying (BPSK) signal constellation in the (M,2;1) AST/STBC and (M,2;2) AST/STBC schemes only. Those studies are far from the exhaustive research.

In other words, the exhaustive research on the spatial diversity order of transmitter diversity ASTs is still missing even for space-time coded systems with coherent detection. For space-time coded systems with non-coherent detection, such as the systems using DSTBCs, the study on the spatial diversity order of AST/DSTBC schemes has not been examined yet. Due to this reason, in this paper, we do not have ambition to examine this issue for all cases, which certainly requires a lot of studies in future.

Instead, we show that the problem of finding the spatial diversity order of the ASTs proposed for channels using DSTBCs with *differential detection* is the same as that problem for the case of *coherent detection* when SNR$\gg$1. Once this has been shown, we consider the (M,2;1) AST/DSTBC and (M,2;2) AST/DSTBC schemes. Since the respective (M,2;1) AST/STBC and (M,2;2) AST/STBC schemes for channels using STBCs provide a full spatial diversity order [10, 29], then the (M,2;1) AST/DSTBC and (M,2;2) AST/DSTBC schemes for channels using DSTBCs also provide a full spatial diversity order as if all Tx and Rx antennas were used.

We restrict ourselves to consider only the *general* $(M,N;K)$ AST/DSTBC scheme for illustration. Other schemes, such as the *restricted* $(M,N;K)$ AST/DSTBC scheme or the $(N+\bar{N},N;K)$ AST/DSTBC scheme are similarly analyzed.

To begin with, we review some crucial discussions mentioned in [10] on the spatial diversity order achieved by the (M,2;K) AST/STBC schemes for channels using STBCs with *coherent detection*. We use the superscript l ($l=1,2,3,\ldots$) to indicate the different coherent durations of the channel. Since the coherent detection is being considered, the channel coefficients between Tx and Rx antennas denoted by $\bar{a}_{ij}^{(l)}$, for $i=1,\ldots,K$ and $j=1,\ldots,M$, are assumed to be perfectly known at the receiver and partially known at the transmitter through a feedback channel. Let $\xi_j^{(l)}=\sum_{i=1}^{K}|\bar{a}_{ij}^{(l)}|^2$. We assume that $\bar{a}_{ij}^{(l)}$s are i.i.d. complex Gaussian random variables with the distribution $\mathcal{CN}(0,\sigma_a)$.

With the notation mentioned in Section 3 of this paper, we rewrite the Tx antenna selection criterion, which was mentioned by Eq. (1) in [10], for the (M,2;K) AST/STBC scheme during the lth coherent duration as

$$\begin{aligned}\hat{\jmath}_2^{(l)} &= \mathcal{F}_2\left(\xi_1^{(l)},\xi_2^{(l)},\ldots,\xi_M^{(l)}\right)\\ &= \mathcal{F}_2\left(\sum_{i=1}^{K}|\bar{a}_{i1}^{(l)}|^2,\sum_{i=1}^{K}|\bar{a}_{i2}^{(l)}|^2,\ldots,\sum_{i=1}^{K}|\bar{a}_{iM}^{(l)}|^2\right). \quad (19)\end{aligned}$$

Denote $\gamma = \frac{E_b}{N_0}$ to be the SNR per bit. It has been shown in [10], the BER expression, say $P_{2,1}$, of the (M,2;1) AST/STBC, where there is only 1Rx antenna, in flat Rayleigh fading channels for binary phase shift keying modulation asymptotically approaches (see Eq. (7) in [10]):

$$P_{2,1} \approx \frac{(2M-1)!}{2^{2M-1}(M-1)!}\left(\frac{1}{\gamma}\right)^M$$

when $\gamma \to \infty$. This equation shows that a full diversity order of M is achieved asymptotically for the (M,2;1) AST/STBC when $\gamma \to \infty$.

The BER expression, say $P_{2,2}$, of the (M,2;2) AST/STBC, where there are 2Tx antennas, in flat Rayleigh fading channels for BPSK modulation asymptotically approaches (see Eq. (8) in [10]):

$$P_{2,2} \approx \frac{M(4M-1)!}{2^{5M-2}(2M-1)(2M-1)!}\left(\frac{1}{\gamma}\right)^{2M}$$

when $\gamma \to \infty$. This equation shows that a full diversity order of $2M$ is achieved asymptotically for the (M,2;2) AST/STBC when $\gamma \to \infty$.

The cases for $K \geq 3$ are not practically significant since it is difficult to employ more than 2Tx antennas at the mobile set in mobile communication downlinks. Due to this reason, the cases for $K \geq 3$ were not presented in [10].

Now we return to consider our proposed, *general* (M,2;K) AST/DSTBC for channels using DSTBCs with *differential detection*. The superscript k ($k = 1,2,3,\ldots,m$) is used to indicate the different coherent durations of the channel (see Fig. 1). Since the differential detection is considered, the channel coefficients between Tx and Rx antennas $a_{ij}^{(k)}$, for $i = 1,\ldots,K,\ j = 1,\ldots,M,\ k = 1,\ldots,m$, are unknown at either the receiver or the transmitter.

As mentioned in Eq. (13) in Section 5, the selection criterion for the *general* (M,2;K) AST/DSTBC during the kth coherent duration is:

$$\begin{aligned}\hat{\mathfrak{I}}_2^{(k)} &= \mathcal{F}_2\left(\chi_1^{(k)},\ldots,\chi_M^{(k)}\right)\\ &= \mathcal{F}_2\left(\sum_{i=1}^{K}|r_{0i1}^{(k)}|^2,\sum_{i=1}^{K}|r_{0i2}^{(k)}|^2,\ldots,\sum_{i=1}^{K}|r_{0iM}^{(k)}|^2\right)\\ &= \mathcal{F}_2\left(\sum_{i=1}^{K}|a_{i1}^{(k)}+n_{0i1}^{(k)}|^2,\sum_{i=1}^{K}|a_{i2}^{(k)}+n_{0i2}^{(k)}|^2,\ldots,\right.\\ &\quad\left.\sum_{i=1}^{K}|a_{iM}^{(k)}+n_{0iM}^{(k)}|^2\right).\end{aligned} \tag{20}$$

We assume that the channel coefficients $a_{ij}^{(k)}$s and noise $n_{0ij}^{(k)}$s are i.i.d. complex Gaussian random variables with the distribution $\mathcal{CN}(0,\sigma_a)$ and $\mathcal{CN}(0,\sigma)$, respectively. We consider the mean and the variance of the following term:

$$\mu_{ij}^{(k)} \triangleq |a_{ij}^{(k)}+n_{0ij}^{(k)}|^2$$

for $i = 1,\ldots,K,\ j = 1,\ldots,M$ and $k = 1,\ldots,m$.

Since $a_{ij}^{(k)}$ and $n_{0ij}^{(k)}$ are the i.i.d. zero-mean, complex Gaussian random variables, $(a_{ij}^{(k)}+n_{0ij}^{(k)})$ are the i.i.d., complex Gaussian random variables with the distribution $\mathcal{CN}(0,\rho)$ where $\rho = \sigma_a+\sigma$. Therefore, $\mu_{ij}^{(k)}$ are the i.i.d, central chi-squared random variables with $n = 2$ degrees of freedom and with the following mean and variance [32, p. 42]:

$$\begin{aligned}E\{\mu_{ij}^{(k)}\} &= n\frac{\rho}{2} = \rho\,,\\ \sigma_{\mu_{ij}^{(k)}} &= 2n\left(\frac{\rho}{2}\right)^2 = \rho^2\,.\end{aligned}$$

We investigate the case in which the channel SNR $\gg 1$. Equivalently, the variances of noise terms $n_{0ij}^{(k)}$s are very small in comparison with the variances of $a_{ij}^{(k)}$s, and therefore, $\rho \approx \sigma_a$. As a result, the means and the variances of $\mu_{ij}^{(k)}$ are approximately:

$$\begin{aligned}E\{\mu_{ij}^{(k)}\} &\approx \sigma_a\\ \sigma_{\mu_{ij}^{(k)}} &\approx \sigma_a^2\,,\end{aligned} \tag{21}$$

when $SNR \gg 1$.

On the other hand, we consider the following term:

$$\theta_{ij}^{(k)} = |a_{ij}^{(k)}|^2$$

for $i = 1,\ldots,K,\ j = 1,\ldots,M$ and $k = 1,\ldots,m$.

Similarly analyzed, $\theta_{ij}^{(k)}$ are the i.i.d, central chi-squared random variables having $n = 2$ degrees of freedom with the following mean and variance [32, p. 42]:

$$\begin{aligned}E\{\theta_{ij}^{(k)}\} &= \sigma_a\,,\\ \sigma_{\mu_{ij}^{(k)}} &= \sigma_a^2\,.\end{aligned} \tag{22}$$

From Eqs. (21) and (22), we realize that, $\mu_{ij}^{(k)}$s and $\theta_{ij}^{(k)}$s have the same statistical properties, i.e., means and variances when $SNR \gg 1$. We can rewrite the antenna selection criterion of the ($M,N;K$) AST/DSTBC in Eq. (20) as

$$\hat{\mathfrak{I}}_2^{(k)} \approx \mathcal{F}_2\left(\sum_{i=1}^{K}|a_{i1}^{(k)}|^2,\sum_{i=1}^{K}|a_{i2}^{(k)}|^2,\ldots,\sum_{i=1}^{K}|a_{iM}^{(k)}|^2\right) \tag{23}$$

when $SNR \gg 1$.

Clearly, the antenna selection criterion for the (M,2;K) AST/DSTBC scheme now tends to be the same as the criterion mentioned in Eq. (19) for the (M,2;K) AST/STBC scheme.

We may conclude that, if the channel $SNR \to \infty$, the behavior of the (M,2;K) AST/DSTBC scheme proposed for channels using DSTBCs with differential detection tends to be the same as that of the (M,2;K) AST/STBC scheme mentioned in literature for channels using STBCs with coherent detection, although the (M,2;K) AST/DSTBC scheme is inferior by 3 dB compared to the (M,2;K) AST/STBC scheme due to the fact that the channel coefficients are not known at either transmitter or receiver. As a result, because the (M,2;1) AST/STBC and (M,2;2) AST/STBC schemes achieve a full spatial diversity [10, 29], then so do the (M,2;1) AST/DSTBC and (M,2;2) AST/DSTBC schemes, provided that the channel SNR is very large.

9. Simulation results

In this section, we run some Monte-Carlo simulations to solidify our proposed AST/DSTBC schemes. We consider a wireless link comprising $K = 1$Rx antenna. The channel SNR is defined to be the ratio between the total average power of the received signals and the average power of noise at the Rx antenna during each STS. Note that the numbers of feedback bits which are required for the *general* $(M,N;K)$ AST/DSTBC and the $(N+\bar{N},N;K)$ AST/DSTBC examined in the simulations are calculated by Eqs. (14) and (15), respectively. The number of feedback bit required for the *restricted* $(M,N;K)$ AST/DSTBC is always 1. In simulations, DSTBCs are modulated by a QPSK signal constellation in simulations.

First, the Alamouti DSTBC in Eq. (1) corresponding to $N = 2$ is simulated. We consider 4 following scenarios:

– Alamouti DSTBC without ASTs;

– Alamouti DSTBC with the *general* (3,2;1) AST/DSTBC (2 feedback bits);

– Alamouti DSTBC with the *restricted* (3,2;1) AST/DSTBC (1 feedback bit);

– Alamouti DSTBC with the (2+1,2;1) AST/DSTBC ($N = 2$, $\bar{N} = 1$, 2 feedback bits).

However, as noted earlier in Remark 13 of Section 6.2, the (2+1,2;1) AST/DSTBC has the same BER performance as the *general* (3,2;1) AST/DSTBC, although the time required to process feedback information in the former is shorter than that in the later. For this reason, we do not need to plot the BER performance of the (2+1,2;1) AST/DSTBC scheme.

Furthermore, in each AST/DSTBC scheme, we examine 2 cases where the feedback error rates are assumed to be 4% and 10%. Transmit antennas in the *restricted* (3,2;1) AST/DSTBC are grouped by the scheme mentioned in Fig. 3b.

Note that it would be better if we can compare the performances here with the performance of a DSTBC without ASTs which provides the same spatial diversity order as the diversity order (equal to 3) provided by the proposed AST/DSTBC schemes, i.e., the *general* (3,2;1) AST/DSTBC, the *restricted* (3,2;1) AST/DSTBC and the (2+1,2;1) AST/DSTBC. This means that we should compare the performance of the Alamouti DSTBC (associated with the proposed ASTs) with that of an order-3 DSTBC (without ASTs). However, while the Alamouti DSTBC has a full rate, it is well known that DSTBCs of an order being greater than 2 with a full rate do not exist. For this reason, it is unfair to compare the Alamouti DSTBC with an order-3 DSTBC, because they have different code rates, and consequently, we do not plot the performance of any order-3 DSTBC in the simulation.

As analyzed earlier, channel coefficients must be constant during at least two adjacent code blocks. If T_c denotes the coherent time of the channel, then it is required that:

- $T_c \geq 4$ STSs for the Alamouti DSTBC without ASTs;
- $T_c \geq 5$ STSs for the Alamouti DSTBC with the *general* (3,2;1) AST/DSTBC, with the *restricted* (3,2;1) AST/DSTBC, or with the (2+1,2;1) AST/DSTBC.

Therefore, to compare fairly the performance of the Alamouti DSTBC with different ASTs, the simulation is run for T_c which is not less than 5 STSs. Example 2 in Section 2.2 is one of such practical scenarios.

The performance of the Alamouti DSTBC with and without ASTs is shown in Fig. 8. It can be seen from Fig. 8 that the proposed ASTs significantly improve the BER performance of the channel. Again, the BER performances of the (2+1,2;1) AST/DSTBC is exactly the same as that of the *general* (3,2;1) AST/DSTBC. The main advantage of the (2+1,2;1) AST/DSTBC over the *general* (3,2;1) AST/DSTBC is that the time required to process feedback information is shortened by 16.67% (see Fig. 7). The SNR reductions [dB] gained by our proposed ASTs to achieve the same BER $= 10^{-3}$ as the Alamouti DSTBC without ASTs are given in Table 1.

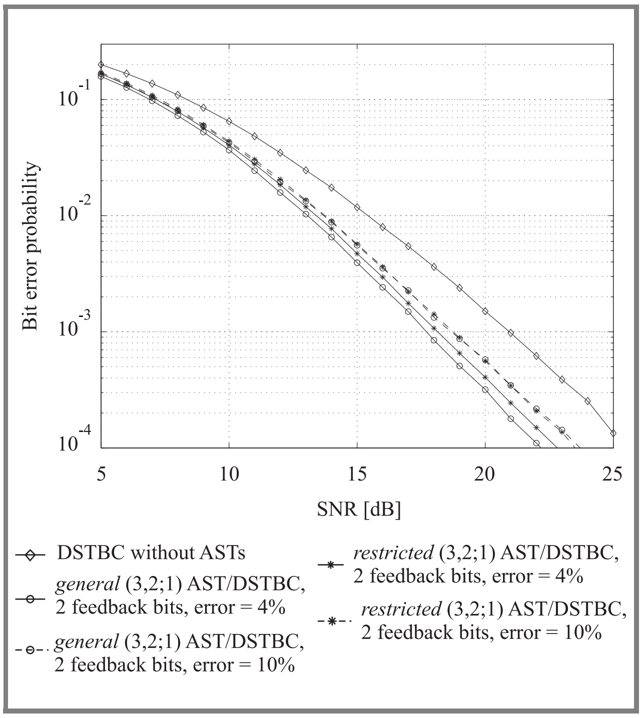


Fig. 8. The Alamouti DSTBC with the *general* (3,2;1) AST/DSTBC and the *restricted* (3,2;1) AST/DSTBC.

Next, we consider the *general* (4,2;1) AST/DSTBC (3 feedback bits) and the *restricted* (4,2;1) AST/DSTBC (1 feedback bit) in which the transmitter selects $N = 2$Tx antennas out of $M = 4$Tx antennas. Clearly, in this case, we have $\bar{N} = M - N = 2$, i.e., $\bar{N} = N$. In Remark 11, we have stated that the (2+2,2;1) AST/DSTBC reduces to the *restricted* (4,2;1) AST/DSTBC. Therefore, we do not plot the performance of the (2+2,2;1) AST/DSTBC here. Transmitter antennas in the *restricted* (4,2;1) AST/DSTBC are grouped by the scheme mentioned in Fig. 3a.

Table 1
SNR reductions [dB] of the *general* (3,2;1) AST/DSTBC, the *restricted* (3,2;1) AST/DSTBC and the (2+1,2;1) AST/DSTBC in the channel using Alamouti DSTBC

Error [%]	*General* (3,2;1) AST/DSTBC	(2+1,2;1) AST/DSTBC	*Restricted* (3,2;1) AST/DSTBC
4	3.25	3.25	2.9
10	2.25	2.25	2.25

Similarly, it is required that:

- $T_c \geq 4$ STSs for the Alamouti DSTBC without ASTs;
- $T_c \geq 6$ STSs for the Alamouti DSTBC with the *general* (4,2;1) AST/DSTBC or with the *restricted* (4,2;1) AST/DSTBC.

To compare fairly the performance of Alamouti DSTBC with different ASTs, the simulation is run for T_c which is not less than 6 STSs. Example 2 mentioned in Section 2.2 is one of such practical scenarios.

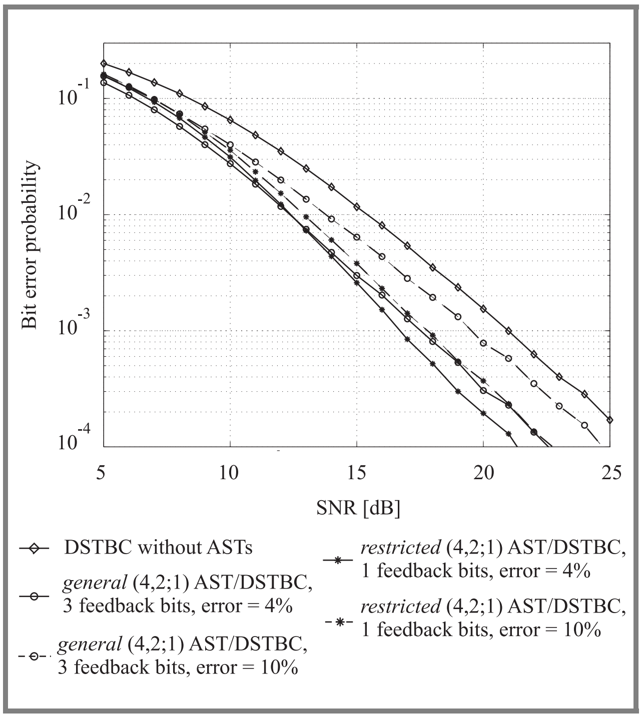


Fig. 9. The Alamouti DSTBC with the *general* (4,2;1) AST/DSTBC and the *restricted* (4,2;1) AST/DSTBC schemes.

The performance of the proposed AST/DSTBC schemes is presented in Fig. 9. The SNR reductions [dB] achieved by our proposed ASTs to have the same BER $= 10^{-3}$ as the DSTBC without ASTs are given in Table 2.

Table 2
SNR reductions [dB] of the *general* (4,2;1) AST/DSTBC and the *restricted* (4,2;1) AST/DSTBC in the channel using Alamouti DSTBC

Error [%]	*General* (4,2;1) AST/DSTBC	*Restricted* (4,2;1) AST/DSTBC
4	3.5	4.3
10	1.5	3.25

Finally, we examine the square, order-4, unitary DSTBC in Eq. (2) corresponding to $N = 4$ and the code rate 3/4. We consider the following 4 scenarios:

– DSTBC without ASTs;

– DSTBC with the *general* (5,4;1) AST/DSTBC (3 feedback bits);

– DSTBC with the *restricted* (5,4;1) AST/DSTBC (1 feedback bit);

– DSTBC with the (4+1,4;1) AST/DSTBC ($N = 4$, $\bar{N} = 1$, 3 feedback bits). Similarly, the BER performance of the (4+1,4;1) AST/DSTBC is exactly the same as that of the *general* (5,4;1) AST/DSTBC, and therefore, we do not need to plot the BER performance of the (4+1,4;1) AST/DSTBC in the simulation.

In each AST, we also consider 2 cases where the feedback error rates are assumed to be 4% and 10%. Transmitter antennas in the *restricted* (5,4;1) AST/DSTBC are grouped by the scheme mentioned in Fig. 3c.
It is required that:

- $T_c \geq 8$ STSs for DSTM without ASTs;
- $T_c \geq 9$ STSs for DSTM with the *general* (5,4;1) AST/DSTBC, with the *restricted* (5,4;1) AST/DSTBC or with the (4+1,4;1) AST/DSTBC.

Therefore, the simulation is run for T_c which is not less than 9 STSs. Example 2 in Section 2.2 is still valid for this scenario.
The performance of the proposed AST/DSTBC schemes is presented in Fig. 10. It is noted that the (4+1,4;1) AST/DSTBC provides the same BER performance as that of the *general* (5,4;1) AST/DSTBC (see Remark 13 in Section 6.2), while shortening the time which is required to process feedback information by 13.33% (see Fig. 7) compared to the *general* (5,4;1) AST/DSTBC.
The SNR reductions [dB] achieved by our proposed ASTs to have the same BER $= 10^{-3}$ as the DSTBC without ASTs are given in Table 3.
From all the above simulations, we realize that the proposed ASTs significantly improve the performance of wireless channels using DSTBCs. Also, we realize that the *restricted*

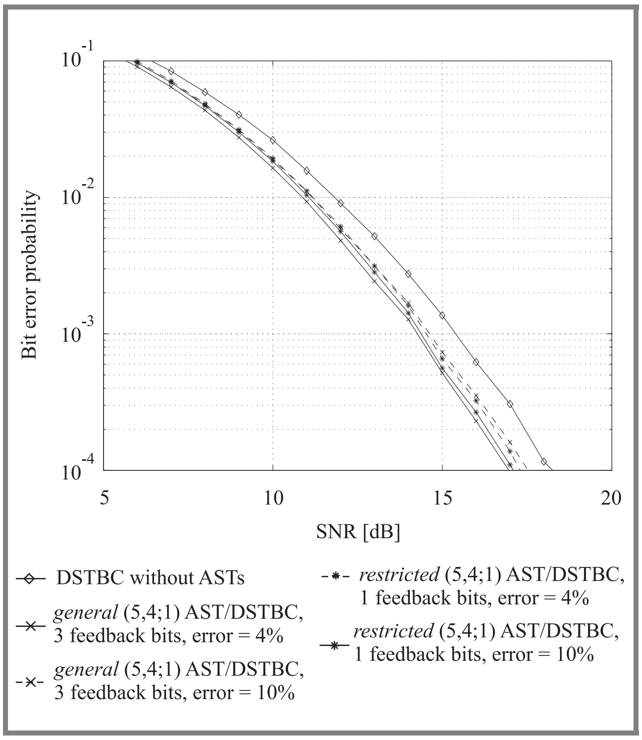


Fig. 10. Square, order-4, unitary DSTBC with the *general* (5,4;1) AST/DSTBC and the *restricted* (5,4;1) AST/DSTBC schemes.

$(M,N;K)$ AST/DSTBC provide a relatively good BER performance compared to the *general* $(M,N;K)$ AST/DSTBC and the $(N+\bar{N},N;K)$ AST/DSTBC, while requiring only 1 feedback bit. More importantly, the *restricted* $(M,N;K)$ AST/DSTBC may perform even better than the *general* $(M,N;K)$ AST/DSTBC and the $(N+\bar{N},N;K)$ AST/DSTBC when the feedback error rate grows large. Intuitively, this

Table 3
SNR reductions [dB] of the proposed (5,4;1) AST/DSTBCs in the channel using square, order-4, unitary DSTBC

Error [%]	*General* (5,4;1) AST/DSTBC	(4+1,4;1) AST/DSTBC	*Restricted* (5,4;1) AST/DSTBC
4	1.2	1.2	1
10	0.8	0.8	0.85

is interpreted by the fact that the *restricted* AST requires only 1 feedback bit while the remaining ASTs require multiple feedback bits. Therefore, when the feedback error rate grows large, the feedback information in the *restricted* ASTs is less likely erroneous than that in the other ASTs. As a result, the *restricted* ASTs are the practical candidates for the channels where fading changes fast.

10. Discussions and conclusion

In this paper, we propose three ASTs referred to as the *general* $(M,N;K)$ AST/DSTBC, the *restricted* $(M,N;K)$ AST/DSTBC, and the $(N+\bar{N},N;K)$ AST/DSTBC for the channels using DSTBCs with arbitrary number of Tx and Rx antennas.

Since the *general* $(M,N;K)$ AST/DSTBC scheme requires a large number of feedback bits when M, N and K are large, it is either impractical or uneconomical for implementation in such cases. The *restricted* $(M,N;K)$ AST/DSTBC and the $(N+\bar{N},N;K)$ AST/DSTBC schemes overcome this shortcoming.

Particularly, the *restricted* $(M,N;K)$ AST/DSTBC is an attractive technique, which provides relatively good bit error performance, compared to the *general* $(M,N;K)$ AST/DSTBC, while requiring only 1 feedback bit. This advantage is very important in the case where the capacity limitation of the feedback channel, such as in the uplink channels of the 3G mobile communication systems, is considered. This advantage is also very beneficial in the channels where fading changes fast and/or the feedback error rate in the feedback channel grows large.

Unlike the restricted AST/DSTBC schemes, where we try to reduce the number of feedback bits, in the $(N+\bar{N},N;K)$ AST/DSTBC schemes, we reduce the average time required to process feedback information. These techniques use at most the same number of feedback bits and provide the same BER performance (if $\bar{N}=1$) as that of the *general* $(M,N;K)$ AST/DSTBC schemes ($M=N+\bar{N}$), but remarkably reduce the average time required to process feedback information.

Simulation show that all three proposed ASTs with a limited number (typically, 1 or 2) of training symbols per each coherent duration of the channel noticeably improve the BER performance of wireless channels utilizing DSTBCs. The improvement is significant even for the case of 1 training symbol, i.e., in the *general* $(M,N;K)$ AST/DSTBC where $M=(N+1)$; in the *restricted* $(M,N;K)$ AST/DSTBC where $M=(N+1)$; or in the $(N+1,N;K)$ AST/DSTBC schemes.

The *restricted* $(M,N;K)$ AST/DSTBC may provide a better BER performance over the *general* $(M,N;K)$ AST/DSTBC and the $(N+\bar{N},N;K)$ AST/DSTBC when the feedback error rate is large. Hence, the restricted AST/DSTBC schemes are a good choice for the channels where fading changes fast and/or the feedback error rate is large.

It is noted that, in this paper, we assume that the carrier phase/frequency is perfectly recovered at the receiver. In fact, phase/frequency recovery errors may exist, which degrade the performance of the proposed ASTs. Those errors may occur due to the difference between the frequency of the local oscillators at the transmitter and the receiver, and/or due to the Doppler frequency-shift effect. The effect of imperfect carrier recovery on the performance of the proposed ASTs in wireless channels utilizing DSTBCs has been examined in our paper [4]. Readers may refer to [4] for more details.

Also, in this paper, the delay of feedback information has not been considered. In reality, the delay of feedback information may somewhat degrade the overall performance of the proposed ASTs. This issue will be mentioned in our other works. Finally, as mentioned earlier, the exhaustive research on the spatial diversity order of the ASTs proposed for channel using DSTBCs has not been derived yet and it must be fully examined in the future work.

References

[1] L. C. Tran, T. A. Wysocki, and A. Mertins, "Improved antenna selection technique to enhance the downlink performance of mobile communications systems", in *Proc. 7th Int. Symp. Sig. Proces. Its Appl. ISSPA 2003*, Paris, France, 2003, vol. 1, pp. 257–260.

[2] L. C. Tran and T. A. Wysocki, "Antenna selection scheme for wireless channels utilizing differential space-time modulation", in *Proc. 7th Int. Symp. Dig. Sig. Proces. Commun. Syst. DSPCS'2003*, Gold Coast, Australia, 2003, pp. 452–457.

[3] L. C. Tran and T. A. Wysocki, "On some antenna selection techniques for wireless channels utilizing differential space-time modulation", in *Proc. IEEE Wirel. Commun. Netw. Conf. WCNC 2004*, Atlanta, USA, 2004, vol. 2, pp. 1210–1215.

[4] L. C. Tran, T. A. Wysocki, J. Seberry, and A. Mertins, "The effect of imperfect carrier recovery on the performance of the diversity antenna selection technique in wireless channels utilizing DSTM", in *Proc. 2nd IEEE Int. Conf. Inform. Technol. Res. Edu. ITRE 2004*, London, UK, 2004, vol. 1, pp. 15–18.

[5] M. Katz, E. Tiirola, and J. Ylitalo, "Combining space-time block coding with diversity antenna selection for improved downlink performance", in *Proc. 54th IEEE Veh. Technol. Conf. VTC 2001*, Atlantic City, USA, 2001, vol. 1, pp. 178–182.

[6] D. A. Gore and A. J. Paulraj, "Space-time block coding with optimal antenna selection", in *Proc. IEEE Int. Conf. Acoust. Speech Sig. Proces. ICASSP '01*, Salt Lake City, USA, 2001, vol. 4, pp. 2441–2444.

[7] T. Xiaofeng, H. Harald, Y. Zhuizhuan, Q. Haiyan, and Z. Ping, "Closed loop space-time block code", in *Proc. IEEE Veh. Technol. Conf. VTC' 2001*, Atlantic City, USA, 2001, vol. 2, pp. 1093–1096.

[8] Z. Chen, B. Vucetic, J. Yuan, and K. L. Lo, "Analysis of transmit antenna selection/ maximal-ratio combining in Rayleigh fading channels", in *Proc. Int. Conf. Commun. Technol. ICCT 2003*, Beijing, China, 2003, vol. 2, pp. 1532–1536.

[9] Z. Chen, "Asymptotic performance of transmit antenna selection with maximal-ratio combining for generalized selection criterion", *IEEE Commun. Lett.*, vol. 8, no. 4, pp. 247–249, 2004.

[10] Z. Chen, J. Yuan, B. Vucetic, and Z. Zhou, "Performance of Alamouti scheme with transmit antenna selection", *Electron. Lett.*, vol. 39, no. 23, pp. 1666–1668, 2003.

[11] Z. Chen, G. Zhu, J. Shen, and Y. Liu, "Differential space-time block codes from amicable orthogonal designs", in *Proc. IEEE Conf. Wirel. Commun. Netw. WCNC 2003*, New Orleans, USA, 2003, vol. 2, pp. 768–772.

[12] G. Ganesan and P. Stoica, "Differential detection based on space-time block codes", *Wirel. Pers. Commun. Int. J.*, vol. 21, May 2002.

[13] G. Ganesan and P. Stoica, "Differential modulation using space-time block codes", *IEEE Sig. Proces. Lett.*, vol. 9, no. 2, pp. 57–60, 2002.

[14] B. L. Hughes, "Differential space-time modulation", *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2567–2578, 2000.

[15] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation", *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041–2052, 2000.

[16] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity", *IEEE J. Select. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, 2000.

[17] B. Vucetic and J. Yuan, *Space-Time Coding*. Hoboken: Wiley, 2003.

[18] S. M. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998.

[19] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: performance results", *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 451–460, 1999.

[20] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time blocks codes from orthogonal designs", *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[21] X.-B. Liang, "Orthogonal designs with maximal rates", *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2468–2503, 2003.

[22] A. V. Geramita and J. Seberry, "Orthogonal designs: quadratic forms and Hadamard matrices", *Lecture Notes in Pure and Applied Mathematics*. New York: Marcel Dekker, 1979, vol. 43.

[23] G. Ganesan, "Designing space-time codes using orthogonal designs", Ph.D. thesis, Uppsala University, Sweden, 2002.

[24] A. Song and X.-G. Xia, "Decision feedback differential detection for differential orthogonal space-time modulation with ASPK signals over frequency-non-selective fading channels", in *Proc. IEEE Int. Conf. Commun. ICC'03*, Anchorage, USA, 2003, vol. 2, pp. 1253–1257.

[25] I. E. Telatar, "Capacity of multi-antenna Gaussian channels", Tech. Rep. AT&T Bell Labs, 1995.

[26] T. S. Rappaport, *Wireless Communication: Principles and Practice*, 2nd ed. Upper Saddle River, London: Prentice Hall, 2002.

[27] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach", *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1650–1656, 2001.

[28] Z. Chen, B. Vucetic, J. Yuan, and Z. Zhou, "Performance analysis of space-time trellis codes with transmit antenna selection in Rayleigh fading channels", in *Proc. IEEE Conf. Wirel. Commun. Netw. WCNC 2004*, Atlanta, USA, 2004, vol. 4, pp. 2456–2462.

[29] Z. Chen, B. Vucetic, and J. Yuan, "Comparison of three closed-loop transmit diversity schemes", in *Proc. 57th IEEE Veh. Technol. Conf. VTC 2003*, Jeju, Korea, 2003, vol. 1, pp. 751–754.

[30] Z. Chen, B. Vucetic, and J. Yuan, "Space-time treliss codes with transmit antenna selection", *Electron. Lett.*, vol. 39, no. 11, pp. 854–855, 2003.

[31] A. Ghrayeb and T. M. Duman, "Performance analysis of MIMO systems with antenna selection over quasi-static fading channels", *IEEE Trans. Veh. Technol.*, vol. 52, no. 2, pp. 281–288, 2003.

[32] J. G. Proakis, *Digital Communications*, 4th ed. Boston: McGraw-Hill, 2001.

Le Chung Tran received the excellent B.Eng. degree with the highest distinction and the M.Eng. degree with the highest distinction in telecommunications engineering from Hanoi University of Communications and Transport, and Hanoi University of Technology, Vietnam, in 1997 and 2000, respectively. From March 2002 to July 2005, he studied towards the Ph.D. degree in telecommunications engineering at the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Australia. He is currently working as Associate Research Fellow at the Telecommunications and Information Technology Research Institute (TITR), School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Australia. He has been working as a lecturer at Hanoi University of Communications and Transport, Vietnam, since September 1997 to date. He has achieved numerous national and overseas awards, including WUS – World University Services (twice), Vietnamese Government's Scholarship, UPA – Wollongong University Postgraduate Award, Wollongong University Tuition Fee Waver, and Alexander von Humboldt (AvH) Research

Fellowship during the undergraduate and postgraduate time. His research interests include transmission diversity techniques, mobile communications, space-time processing, MIMO systems, channel propagation modelling, ultra-wide band communications, OFDM, and spread spectrum techniques.

e-mail: lct71@uow.edu.au
University of Wollongong
Northfields Avenue
Wollongong, NSW 2522, Australia

Tadeusz Antoni Wysocki received the M.Sc.E. degree with the highest distinction in telecommunications from the Academy of Technology and Agriculture, Bydgoszcz, Poland, in 1981. In 1984, he received his Ph.D. degree, and in 1990, was awarded a D.Sc. degree (habilitation) in telecommunications from the Warsaw University of Technology. In 1992, Doctor Wysocki moved to Perth, Western Australia to work at Edith Cowan University. He spent the whole 1993 at the University of Hagen, Germany, within the framework of Alexander von Humboldt Research Fellowship. After returning to Australia, he was appointed a Program Leader, Wireless Systems, within Cooperative Research Centre for Broadband Telecommunications and Networking. Since December 1998 he has been working as an Associate Professor at the University of Wollongong, NSW, within the School of Electrical, Computer and Telecommunications Engineering. The main areas of Doctor Wysocki's research interests include: indoor propagation of microwaves, code division multiple access (CDMA), space-time coding and MIMO systems, as well as mobile data protocols including those for ad hoc networks. He is the author or co-author of four books, over 150 research publications and nine patents. He is a Senior Member of IEEE.

wysocki@uow.edu.au
University of Wollongong
Northfileds Ave
Wollongong, NSW 2522, Australia

Alfred Mertins received his Dipl.-Ing. degree from the University of Paderborn, Germany, in 1984, the Dr.-Ing. degree in electrical engineering and the Dr.-Ing. habil. degree in telecommunications from the Hamburg University of Technology, Germany, in 1991 and 1994, respectively. From 1986 to 1991 he was with the Hamburg University of Technology, Germany, from 1991 to 1995 with the Microelectronics Applications Center Hamburg, Germany, from 1996 to 1997 with the University of Kiel, Germany, from 1997 to 1998 with the University of Western Australia, and from 1998 to 2003 with the University of Wollongong, Australia. In April 2003, he joined the University of Oldenburg, Germany, where he is a Professor in the Faculty of Mathematics and Science. His research interests include speech, audio, image and video processing, wavelets and filter banks, and digital communications.

e-mail: alfred.mertins@uni-oldenburg.de
University of Oldenburg
Signal Processing Group
Institute of Physics
26111 Oldenburg, Germany

Jennifer Seberry graduated Ph.D. in computation mathematics from La Trobe University in 1971. She has subsequently held positions at the Australian National University, The University of Sydney and ADFA, The University of New South Wales. She has published extensively in discrete mathematics and is world renown for her new discoveries on Hadamard matrices and statistical designs. In 1970 she co-founded the series of conferences known as the 20th Australian Conference on Combinatorial Mathematics and Combinatorial Computing. She started teaching in cryptology and computer security in 1980. She is especially interested in authentication and privacy. In 1987, at University College, ADFA, she founded the Centre for Computer and Communications Security Research to be a reservoir of expertise for the Australian community. Her studies of the application of discrete mathematics and combinatorial computing via bent functions, S-box design, has led to the design of secure crypto-algorithms and strong hashing algorithms for secure and reliable information transfer in networks and telecommunications. Her studies of Hadamard matrices and orthogonal designs are applied in CDMA technologies. In 1990 she founded the AUSCRYPT/ASIACRYPT series of International Cryptologic Conferences in the Asia/Oceania area. She has supervised 25 successful Ph.D. candidates, has over 350 scholarly papers and six books.

e-mail: jennie@uow.edu.au
University of Wollongong
Northfields Avenue
Wollongong, NSW 2522, Australia

Regular paper

Modeling spatial aspects of mobile channel for low antenna height environments

Mohammed T. Simsim, Noor M. Khan, and Predrag B. Rapajic

Abstract— It is essential to have deep understanding of the mobile radio channel in particular for radio communication modeling and advanced technology system design. Models for the mobile radio channel are vital for the study of smart antenna systems, both for the design of algorithms, system-testing purposes, and for network planning. This paper provides an intensive study of the spatial characteristics of the mobile channel for low antenna height cellular environments, i.e., picocells and microcells, assuming Gaussian distributed scatterers. We investigate previous work on the angle of arrival (AoA) statistics for Gaussian distributed scatterers and make appropriate comments. Further, we employ the recently proposed eccentro-scattering physical channel model, as a generalized model, to derive the probability density function (pdf) of AoA of the multipaths at base station (BS) assuming Gaussian distributed scatterers around both BS and mobile station (MS). We found that the pdf of AoA at BS is directly affected by the standard deviation of the scatterers' density and the size of the scattering disc. The derived formulas, in closed form, can be used further for performance assessment of smart antennas and beamwidth design purposes.

Keywords— Gaussian distributions, geometric modeling, multipath channels, scattering.

1. Introduction

During the last few years, there has been a great interest in the research to enhance the capacity of the mobile radio channel to meet the increasing demand of large data rate transmissions. Towards this goal, more attention was given to the spatial domain of the radio propagation channel in order to employ it more efficiently. While conventional transceivers use temporal and spectral signal processing, modern transceivers need to exploit the spatial dimension, as well. Smart antenna technology was proposed as a promising technical solution that utilizes the spatial parameters of the radio propagation channel to enhance the performance [1, 2].

In order to exploit the spatial dimension efficiently, reliable channel models are required which would lead to the design of effective signal processing schemes. Wireless signal has unpredictable behavior due to multipath fading, which confronts the effective exploitation of the mobile radio channel. Complete knowledge of the multipaths would bring about great improvements in wireless communication services in terms of capacity and performance.

Radio environments have extremely different geographical and electrical features, which in turn lead to different propagation mechanisms. However, similar propagation mechanisms have been characterized into three main categories based on the cell type, i.e., picocell, microcell, and macrocell environments. Yet, there would still be some variability within these categories due to differences in antenna heights of base station (BS) and mobile station (MS), average height of the surrounding buildings, distance between BS and MS, total number of users, distances between users, etc.

The idea of generic channel modeling and the description of its parameters based on some measurements has been presented in [3, 4]. It may take few more years to obtain the correct parameters for such generic models from rigorous measurement campaigns [4]. Even though, many geometry-based physical channel models have been proposed in the literature [5–15] but still the discussion is open to adopt set of values as the actual parameters for different cellular environments [4].

Previously proposed geometry-based physical channel models are specific for modeling particular cellular environments. For instance, circular scattering models (CSM) were proposed to model the scattering environment in macrocells [9–11] and elliptical scattering model (ESM) was proposed to model the scattering environment in picocells and microcells [11, 12]. The distribution of scatterers within the scattering area is considered to be either uniform or Gaussian. In uniformly distributed scattering region, scatterers are assumed to have constant density throughout the scattering area, which simplifies the analysis and manipulation. Whereas for Gaussian distributed scatterers, the majority of scattering points are situated close to MS and the density of scattering points decreases as the distance from MS increases. This is a more practical approach to model real situations.

Considering the proximity of a model to a particular environment, the previous models are very specific to their environments either with respect to the shape of the scattering disc [9–12] or with respect to the standard deviation of the Gaussian distributed scatterers around MS [13]. An elliptical model was proposed in [14], which could be used as a circular one as well with change in eccentricity of the ellipse, for uniformly distributed scatterers, only. Even though, the model was able to explain the radio wave propagation phenomenon in macrocell and quasi-macrocell environments but did not deliver enough information to model picocell and microcell environments where both MS and BS fall inside the scattering region. This is due to the fact that, in picocell and microcell environments, the antenna heights, of transmitter and receiver, are lower as compared to those in macrocell and quasi-macrocell environments. Therefore, multipath scattering is assumed near BS and MS.

Gaussian scatter density model (GSDM) was proposed in [13] to be applied to every type of cellular environment by changing the standard deviation of the scatterers around MS, only. Thus, to model a picocell environment using GSDM, the standard deviation of the scatterers around MS is increased in order to encompass BS inside the scattering area. This way, BS is supposed to be inside the scattering region but it falls at its edge, hence the density of scatterers around BS would be much less than that around MS, which is not always the case especially in picocell environments. Usually in picocells, scatterers exist in the vicinity of BS as much as around MS. Moreover, a comparison is presented in [13] among GSDM, CSM, and ESM to show the superiority of GSDM. It is worth noting that GSDM assumes Gaussian distribution for the scatterers around MS without confining the scatterers in a scattering disc while CSM and ESM assume uniform distribution of the scatterers confined in scattering discs, which may include MS only (for CSM), or both MS and BS (for ESM). Therefore, this comparison makes no clear distinction between the shape of the scattering disc and the distribution of scatterers within that disc.

A generalized physical channel model has been recently proposed in [16] that can be used to model all types of cellular environments, i.e., picocell, microcell, macrocell, uniform scattering, Gaussian scattering, bounded scatterers, unbounded scatterers with appropriate choice of few model parameters. This provides a unified approach to model all cellular environments. The proposed eccentro-scattering model can be easily used to derive the probability density function (pdf) of angle of arrival (AoA) of the multipaths at BS and/or MS.

Here, we employ the eccentro-scattering physical channel model [16] to derive the pdf of AoA of the multipaths at BS from Gaussian distributed scattering regions around MS and BS for picocell and microcell environments while we carry similar work for macrocell environment in another paper. The derived formulas for the pdf of AoA in closed form are useful for analytical and simulation purposes.

The angular distribution of the received multipath components is very useful in systems employing spatial filtering, e.g., smart antennas, to determine the performance of radio link [12]. Even thought, it is possible to obtain the pdf of AoA from measured data or from site-specific propagation prediction techniques, however this type of data may not be always available. Therefore, statistical models are helpful to characterize the AoA of the multipaths. Furthermore, being a generalized model yields more flexibility, since it is easy to imitate several environments/situations by changing the model parameters.

Angle of arrival statistics are helpful in the performance and capacity enhancement of multiple input multiple output (MIMO) systems. Spatial division multiple access (SDMA) is a special case of MIMO systems intended for multiuser detection in which BS antenna coefficients are optimized, adaptively, for radiating maximum energy towards the desired user keeping it orthogonal to all other users in spatial domain. So, the pdf of AoA becomes an important factor in designing beamwidth and incidence angle for SDMA case.

Our contributions in this paper are as follows:

- We inspect previous work [9, 13, 15, 17] on the scatterers' density and AoA statistics for Gaussian distributed scatterers and make appropriate comments and comparisons.
- We exploit the recently proposed eccentro-scattering physical channel model to derive the pdf of AoA of the multipath signals at BS in closed form for Gaussian distributed scatterers around both BS and MS. That is, the model assumes Gaussian distributed scatterers around MS in addition to another Gaussian distributed scatterers around BS, each with different standard deviation, in order to provide more flexibility in the design and to simulate the real situations more closely. Being able to independently change the standard deviations of these two Gaussian functions means that we can have a model of Gaussian distributed scatterers around MS only, e.g., GSDM, or a model of Gaussian distributed scatterers around BS only as special cases. Whereas the general case corresponds to the model of Gaussian distributed scatterers around both BS and MS where increasing (or decreasing) the corresponding standard deviation indicates less (or more) scattering density in the vicinity of the respective antenna.
- We consider Gaussian distributed scatterers confined in a scattering disc and investigate the advantages of this technique.

The rest of the paper is organized as follows: in Section 2 we present the physical channel model and describe its parameters, Section 3 presents the derivation of the pdf of AoA of the multipaths at BS and Section 4 concludes the paper.

2. Channel model

2.1. General remarks

Wave propagation path changes in different environments according to scatterers' density and distance between BS and MS depending on the maximum delay spread. The propagated radio signal experiences shorter delays in picocell and microcell environments as compared to macrocell environment. The word "scattering" is not only

used for diffuse scattering but even for those processes that are strictly speaking "'specular reflections" [3].
In picocell environment, BS and MS are few meters away from each other and are surrounded by local scatterers (Fig. 1a). The antenna heights are relatively low and multipath scattering is assumed near BS as likely as around MS. This situation mainly occurs in indoor wireless communication, i.e., offices and factory/hall, and it may include street crossings under some circumstances.

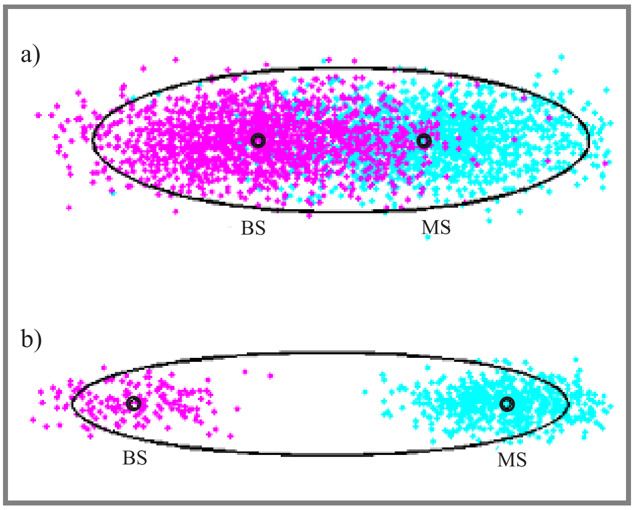


Fig. 1. Picocell (a) and microcell (b) environments.

In microcell environment, the distance between BS and MS is greater than that in picocell environment (typically in the order of a few hundred meters [3]). Here also antenna heights are relatively low, however BS usually has less scattering points in its vicinity as compared to MS (Fig. 1b). This situation corresponds, in general, to streets and open areas such as shopping malls, busy roads, and downtown areas.
Our study is based on the following assumptions.

- Scatterers are enclosed in an elliptical scattering disc whose eccentricity can be altered according to the maximum delay spread and the distance between BS and MS. Considering unbounded scattering regions may simplify the derivation but it is not practical since multipaths with longer delays experience greater path loss and therefore have relatively low power as compared to those with shorter delays [12, 17]. The scattering disc, which is referred to as "the eccentro-scattering disc" [16], can simulate both circular and elliptical models with corresponding choice of eccentricity.

- Scatterers are Gaussian distributed within the scattering disc. We investigate the relation between the standard deviation of the distribution of scatterers and the pdf of AoA of the multipath signals at BS.

- Multipath signals received at the antenna are plane waves coming from the horizon, i.e., only azimuthal coordinate is considered.

- Received signal at the antenna undergoes no more than one reflection by scatterers when traveling between transmitter and receiver. Placing a scattering object at the last re-radiation and approximating the preceding scattering as a stochastic process can retain some of the properties of multiple bounces while providing for a much simpler model [18]. Practically, we are considering only the distribution of scatterers contributing to the last reradiation while the preceding multiple-bounce can be modeled as a stochastic process which has lognormal shadowing with Nakagami fading [19]. This assumption, in conjunction with the first assumption, means that we are considering all scatterers giving rise to a single bounce multipath signal arriving at the receiving antenna up to $\tau_{\max}$. Hence, $\tau_{\max}$ is the maximum allowed delay spread, i.e., the time difference between first and last signal arrivals at the receiving antenna with signal power exceeding some threshold value defined by the system designer.

- Each scatterer is assumed to be an omnidirectional re-radiating element with equal scattering coefficients and uniform random phases.

- Effective antenna pattern is omnidirectional. Practically, the derived formulas for the pdf of AoA should be used in conjunction with the actual antenna radiation pattern.

2.2. Model description

The elliptical diagram shown in Fig. 2 represents the eccentro-scattering model for picocell and microcell environments. In the figure, C is the center of the ellipse with foci B (BS) and M (MS) separated by a distance D and with semi-major and semi-minor axes a and b, respectively. The line segments r_b and r_m satisfy

$$r_b + r_m = 2a\,. \tag{1}$$

Local scatterers exist in the vicinity of BS and MS in picocell and microcell environments, as mentioned before. Therefore both BS and MS are placed inside the elliptical scattering disc. Based on the eccentro-scattering model,

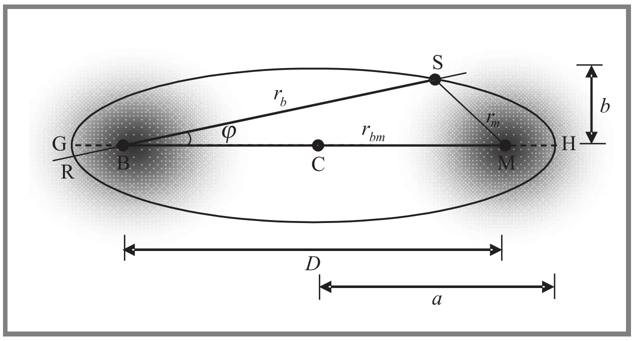


Fig. 2. Eccentro-scattering model for picocells and microcells.

BS and MS are located at the focal points of the eccentro-scattering disc.

Exemplary values of the model parameters are listed in Table 1. The values in the table are merely typical ones while more specific values may differ greatly from those on the table. For example, the delay spread in corridor (picocells) may exceed 0.1 μs in some cases due to the variety of antennas, measurement techniques, measurement locations, and different assumptions made by the working groups. However, we have based our numbers on values given in [1, 3]. The standard deviations of the Gaussian distributed scatterers have been taken from [13, 17].

Table 1
Typical values of the parameters of eccentro-scattering model for picocell and microcell environments

Environment	Picocell	Microcell
Maximum delay spread ($\tau_{\max}$) [μs]	≤ 0.1	0.3
Angle spread ($2\phi_{\max}$) [degrees]	360	120
Distance between BS and MS (D) [m]	$1 < D < 30$	$50 < D < 500$
Semi-major axis (a) [m]	$15 < a < 65$	$45 < a < 545$
Eccentricity (e)	$0 < e < 1$	$0.33 < e < 1$
Standard deviations (σ_b, σ_m) of scatterers around BS and MS	$\sigma_b = \sigma_m$ $ae \leq \sigma_m \leq 2ae$	$\sigma_b \geq \sigma_m$ $0.4ae \leq \sigma_m \leq 0.68ae$

Using radio signal propagation theory, the relationship among maximum delay spread, $\tau_{\max}$, total path traveled, $r_b + r_m$, and speed of light, c, can be written as

$$r_b + r_m = c\tau_{\max}\,. \tag{2}$$

From Eqs. (1) and (2), the semi-major axis of the eccentro-scattering disc, a, in Fig. 2 is

$$a = \frac{c\tau_{\max}}{2}\,. \tag{3}$$

The eccentricity, e, aspect ratio, k, and semi-minor axis, b, of the eccentro-scattering disc in Fig. 2 are:

$$e = \frac{D}{2a}\,, \tag{4}$$

$$k = \frac{b}{a} = \sqrt{1-e^2}\,, \tag{5}$$

$$b = \frac{1}{2}\sqrt{4a^2 - D^2}\,. \tag{6}$$

A Gaussian model of the spatial pdf of the scatterers around MS and BS can be written as in Eqs. (7) and (8), respectively,

$$p_{X_s,Y_s}(x_S, y_S) = \frac{1}{2\pi\sigma_m^2}\exp\left[-\frac{\left|(x_S-x_M)^2+(y_S-y_M)^2\right|}{2\sigma_m^2}\right], \tag{7}$$

$$p_{X_s,Y_s}(x_S, y_S) = \frac{1}{2\pi\sigma_b^2}\exp\left[-\frac{\left|x_S^2+y_S^2\right|}{2\sigma_b^2}\right]. \tag{8}$$

Or in polar coordinates, Eqs. (7) and (8) can be written as in Eqs. (9) and (10), respectively [20],

$$p_{R_b,\Phi_m}(r_b,\varphi) = \begin{cases} \dfrac{\|r_b\|}{2\pi\sigma_m^2}\exp\left[-\dfrac{\|r_b - r_{bm}\|^2}{2\sigma_m^2}\right], & r_b > 0 \\ 0, & \text{elsewhere} \end{cases} \tag{9}$$

$$p_{R_b,\Phi_b}(r_b,\varphi) = \begin{cases} \dfrac{\|r_b\|}{2\pi\sigma_b^2}\exp\left[-\dfrac{\|r_b\|^2}{2\sigma_b^2}\right], & r_b > 0 \\ 0, & \text{elsewhere} \end{cases} \tag{10}$$

where σ_m and σ_b are the standard deviations of the distribution of scatterers around MS and BS, respectively. The angle φ is the angle with the horizontal (in our case the AoA of the multipaths at BS from a scattering point S), and r_b and r_{bm} are the position vectors of the scattering point S and MS, respectively, with respect to BS, as shown in Fig. 2. The angle 0° corresponds to the angle towards MS, i.e., direction of line of sight (LoS). The spatial pdf of the scatterers around MS in polar coordinates presented in [9, 15] is deficient in r_b factor, which is the Jacobian of the transformation from Cartesian coordinate system to polar coordinate system.

3. Probability density function of the angle of arrival

Considering Eq. (9) and the geometry in Fig. 2, the density of scattering points around MS can be described by the bivariate Gaussian distribution as follows:

$$p_{R_b,\Phi_m}(r_b,\varphi) = \begin{cases} \dfrac{r_b}{2\pi\sigma_m^2}\exp\left[-\dfrac{r_b^2+D^2-2r_bD\cos\varphi}{2\sigma_m^2}\right], & r_b > 0 \\ 0, & \text{elsewhere} \end{cases}. \tag{11}$$

In a similar way, the density of scattering points around BS can be defined as

$$p_{R_b,\Phi_b}(r_b,\varphi) = \begin{cases} \dfrac{r_b}{2\pi\sigma_b^2}\exp\left[-\dfrac{r_b^2}{2\sigma_b^2}\right], & r_b > 0 \\ 0, & \text{elsewhere} \end{cases}. \tag{12}$$

For microcell environment, the maximum AoA, $\varphi_{\max}$, has been found to be 60° [1] and the density of scatterers around MS is greater than that around BS. Whereas, for picocell environment, the maximum AoA,$\varphi_{\max}$,, is 180° [1] and the density of scatterers around MS is almost equal to that around BS. The distance between BS and MS, D, is larger in case of microcells than picocells.

The area bounded by the sector BSH in Fig. 2 is a function of the angle φ, with angles between 0 and $\varphi_{\max}$.

Considering Eqs. (11) and (12) and the geometry in Fig. 2, the cumulative distribution functions (CDF) of the scattering points around MS and BS would be respectively defined as

$$P_{\Phi_m}(\varphi) = \int_{-\varphi_{\max}}^{\varphi} \int_{0}^{r_{b1}} \frac{r_b}{2\pi\sigma_m^2} \exp\left(\frac{-r_b^2 - D^2 + 2r_b D\cos\eta}{2\sigma_m^2}\right) dr_b\, d\eta \tag{13}$$

and

$$P_{\Phi_b}(\varphi) = \int_{-\varphi_{\max}}^{\varphi} \int_{0}^{r_{b1}} \frac{r_b}{2\pi\sigma_b^2} \exp\left(\frac{-r_b^2}{2\sigma_b^2}\right) dr_b\, d\eta\,, \tag{14}$$

where r_b is the positive root of the equation defining the disc in Fig. 2 in polar coordinates, i.e.,

$$r_b^2\left(\frac{\cos^2\varphi}{a^2} + \frac{\sin^2\varphi}{b^2}\right) - r_b\left(\frac{D\cos\varphi}{a^2}\right) + \frac{D^2}{4a^2} = 1\,. \tag{15}$$

Solving Eq. (15) for r_b results in the following two solutions:

$$r_{b1,b2} = \frac{4a^2 - D^2}{2(\pm 2a - D\cos\varphi)}\,. \tag{16}$$

Since the received signal at the antenna has interacted with only one single scatterer in the channel, as assumed earlier, then the AoA of the multipaths from scatterers around BS and MS are two disjoint events. Hence, the pdf of AoA of the multipaths from all scattering points within the eccentro-scattering disc, $p_\Phi(\varphi)$, would be basically the addition of the derivatives of Eqs. (13) and (14) with respect to φ, i.e.,

$$\begin{aligned} p_\phi(\varphi) &= \frac{1}{2}\Big(p_{\Phi_m}(\varphi) + p_{\Phi_b}(\varphi)\Big) \\ &= \frac{1}{2}\left(\frac{d}{d\varphi}P_{\Phi_m}(\varphi) + \frac{d}{d\varphi}P_{\Phi_b}(\varphi)\right), \end{aligned} \tag{17}$$

therefore

$$\begin{aligned} p_\Phi(\varphi) &= \int_0^{r_{b1}} \frac{r_b}{2\pi\sigma_m^2} \exp\left(\frac{-r_b^2 - D^2 + 2r_b D\cos\varphi}{2\sigma_m^2}\right) dr_b \\ &+ \int_0^{r_{b1}} \frac{r_b}{2\pi\sigma_b^2} \exp\left(\frac{-r_b^2}{2\sigma_b^2}\right) dr_b\,. \end{aligned} \tag{18}$$

Substituting the value of r_{b1} from Eq. (16) into Eq. (18), we get the pdf of AoA of the multipaths at BS from Gaussian distributed scatterers around BS and MS confined in an eccentro-scattering disc as in Eq. (19):

where Δ is a normalizing constant such that $\int_0^{2\pi} p_\Phi(\varphi)d\varphi = 1$, and $\mathrm{erf}(x)$ is the well known error function defined as $\mathrm{erf}(x) = \int_0^x \exp(-t^2)dt$.

The work in [15] derived the pdf of AoA of the multipath signals at BS considering scattering points that are Gaussian distributed around MS, only, and within the angular beamwidth of a directional antenna. But no geometrical shape of the scattering disc was defined. The only effect of using directional antenna at BS is to reject AoAs falling outside the beamwidth, while it does not alter the distribution of AoA at BS.

As mentioned earlier, maximum delay spread of the multipaths, $\tau_{\max}$, has a significant effect on the pdf of AoA at BS. Figures 3 and 4 show that the pdf of AoA at BS derived using the eccentro-scattering model depends

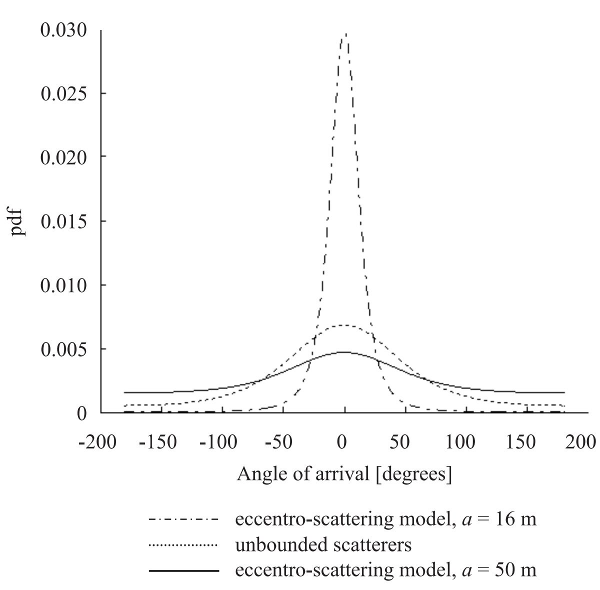


Fig. 3. Probability density function of AoA for picocell environment with $D = 30$ m, $\sigma_b = \sigma_m = D/2$ using eccentro-scattering model and unbounded scattering region.

on the size of the scattering disc, which in turn depends on $\tau_{\max}$, while this is not the case for the corresponding pdf derived using unbounded scatterers [13, 15, 17]. It was mentioned in [13] that "in the Gaussian AoA model, the distribution of arriving waves in azimuth is assumed to be Gaussian without specific mention of the scatter density required to produce it". But it was already stated earlier in [21] that "if the bell-shaped spatial Gaussian model

$$p_\Phi(\varphi) = \frac{\Delta}{4\pi}\left\{\begin{aligned} &1 + \exp\left(-\frac{2a^2e^2}{\sigma_m^2}\right) - \exp\left(-\frac{a^2(1+e^2-2e\cos\varphi)^2}{2\sigma_m^2(1-e\cos\varphi)^2}\right) + \frac{\sqrt{2\pi}ae\cos\varphi}{\sigma_m}\exp\left(-\frac{2a^2e^2\sin^2\varphi}{\sigma_m^2}\right) \\ &\left[\mathrm{erf}\left(\frac{\sqrt{2}ae\cos\varphi}{\sigma_m}\right) + \mathrm{erf}\left(\frac{a(1-2e\cos\varphi+e^2\cos 2\varphi)}{\sqrt{2}\sigma_m(1-e\cos\varphi)}\right)\right] - \exp\left(-\frac{a^2(1-e^2)^2}{2\sigma_b^2(1-e\cos\varphi)^2}\right) \end{aligned}\right\}, \tag{19}$$

for the distribution of scatterers round the mobile is assumed then the angular distribution seen form BS is also Gaussian". In fact, if we consider unbounded scatterers or multipaths from infinitely far distant scatterers, we get exactly bell-shaped Gaussian pdf of AoA of the multipath signals at BS. But if we consider Gaussian distributed scatterers confined in some disc, e.g., eccentro-scattering disc, still the pdf of AoA would be Gaussian distributed while its tails depend on the size of the disc, the distance between BS and MS, and the standard deviation of the scattering points around MS and BS under the zooming effect (Figs. 3 and 4).

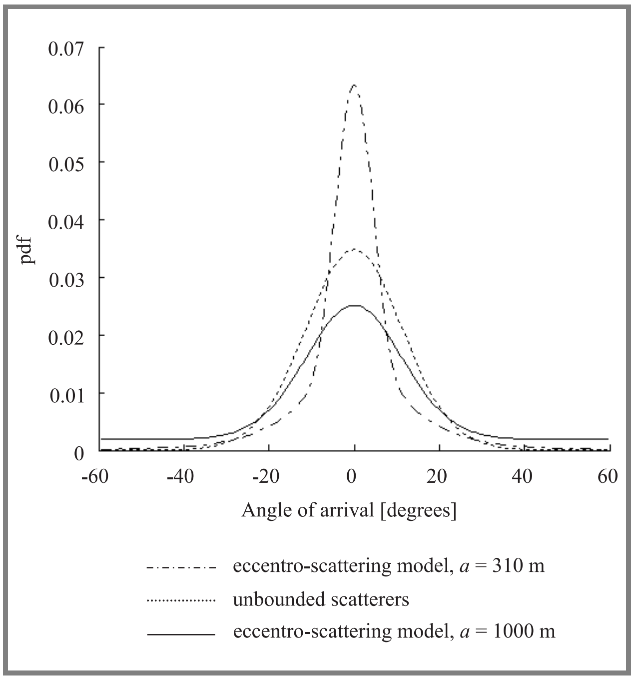


Fig. 4. Probability density function of AoA for microcell environment with $D = 600$ m, $\sigma_b = 0.34D$, $\sigma_m = 0.2D$ using eccentro-scattering model and unbounded scattering region.

If we consider $\sigma_m = \infty$ and $\sigma_b = 0$ in Eq. (19), the pdf of AoA of the multipath signals at BS approaches a uniform distribution. Whereas, if we consider $a = 4 \times \sigma_m$ and $\sigma_b = 0$, Eq. (19) approaches the result found by Janaswamy [13] for unbounded Gaussian distributed scatterers around MS only. This is because at $a = 4 \times \sigma_m$, the eccentro-scattering disc would confine almost all (99.99%) of the scattering points. Whereas, practically, a must be smaller than $4 \times \sigma_m$.

As shown in Fig. 5, the pdf of AoA of the multipath signals at BS derived using the eccentro-scattering model depends on the value of a provided that $a \leq 4 \times \sigma_m$. Therefore, it is more realistic to bound the scatterers inside some scattering disc according to terrain conditions, i.e., eccentro-scattering disc.

Similarly, if we consider $\sigma_m \geq 5 \times a$ and $\sigma_b = 0$, Eq. (19) approaches the results found originally by Liberti [12] and derived again in a compact form by Ertel [11] for bounded uniformly distributed scatterers confined in an elliptical scattering disc.

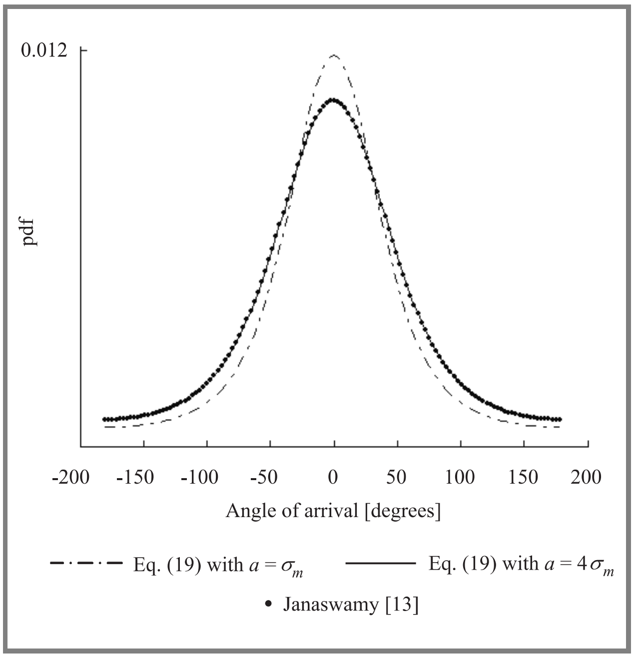


Fig. 5. Effect of increasing a on the pdf of AoA, $\sigma_b = 0$.

Figure 6 shows the effect of using large values for σ_m on the pdf of AoA of the multipaths at BS. We note that Gaussian distributed scatterers confined in an eccentro-scattering disc with large values of σ_m approach uniformly distributed scatterers confined in the disc.

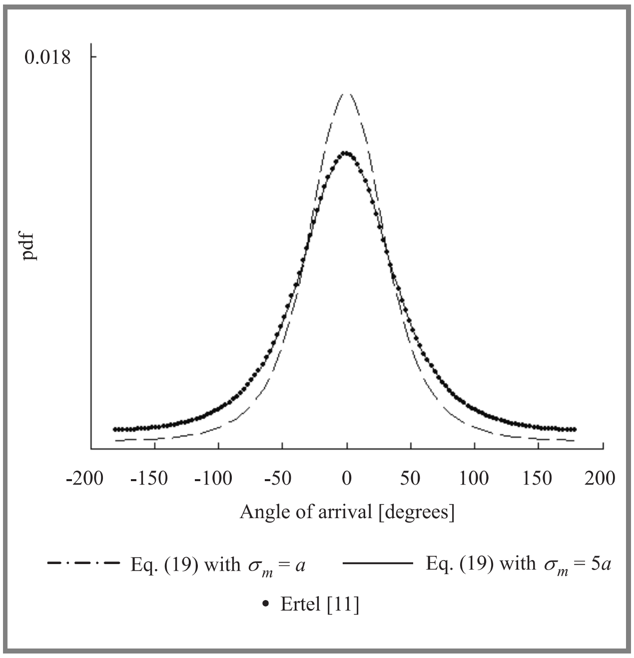


Fig. 6. Effect of increasing σ_m on the pdf of AoA, $\sigma_b = 0$.

In [22], Lapalcian distribution of angular variance 25.5° was proposed as the best fit for indoor measurements, but it was observed that the tails were actually heavier than predicted by Laplacian distribution. The model proposed by Spencer captured the behavior of 90% of arrivals while the eccentro-scattering model represents a better fit (more than 98%). The tail behavior predicted in the proposed models and that observed in the measured data are examined by looking at the distribution on log scale as shown in Fig. 7. Furthermore, the Laplacian function lacks a theoretical explanation [23].

In Figs. 8 and 9, simulation and theoretical results are presented using different values of σ_b and σ_m for typical picocell and microcell environments. In picocell environment, the effect of the standard deviation of scatterers is not so evident because BS and MS are usually located very close to each other while in microcell environment this effect is obvious because BS and MS are relatively far from each other and the scattering disc gets larger values of eccentricity, e, close to 1.

These results illustrate the accuracy and generality of the eccentro-scattering model.

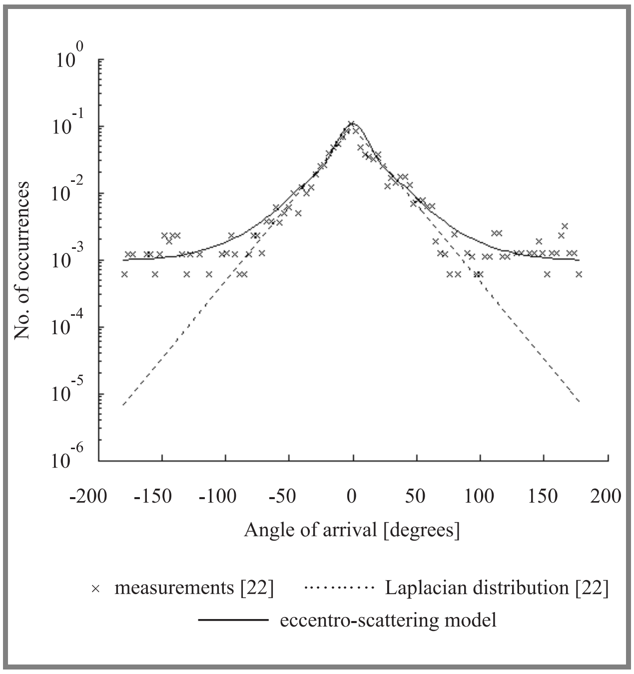


Fig. 7. Comparison of the pdf of AoA for eccentro-scattering model ($a = 36$ m, $D = 64$ m, and $\sigma_m = \sigma_b = 18$), Laplacian distribution ($\sigma = 25.5°$) [22], and measurements [22].

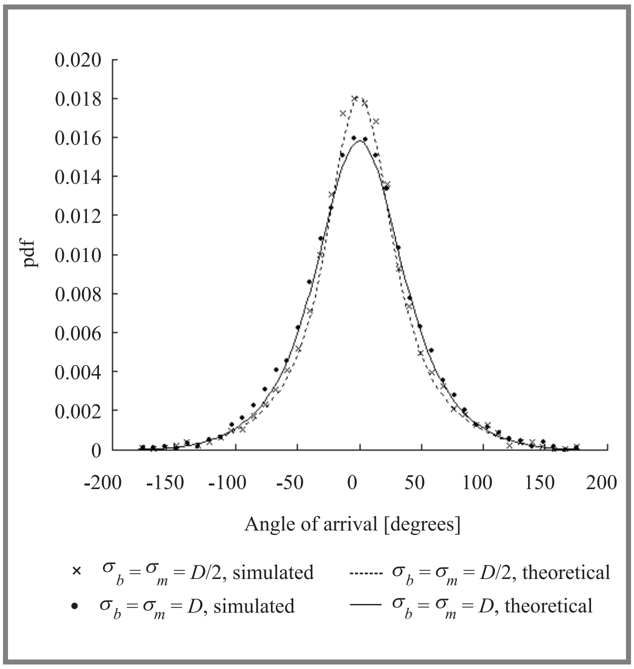


Fig. 8. Simulated and theoretical pdf of AoA for picocells with $D = 30$ m, $a = 25$ m using different values for $\sigma_b = \sigma_m$.

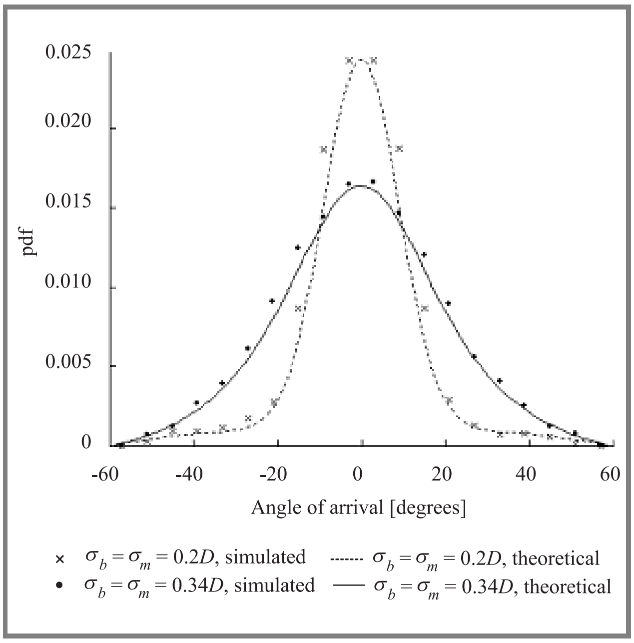


Fig. 9. Simulated and theoretical pdf of AoA for microcells with $D = 300$ m, $a = 180$ m using different values for $\sigma_b = \sigma_m$.

4. Conclusions

In this paper we have presented an intensive study of the spatial statistics of the mobile channel for picocell and microcell environments assuming Gaussian distributed scatterers around both base station (BS) and mobile station (MS) and have made appropriate comments on previous work. We have made clear distinction between the shape of the scattering disc and the distribution of scatterers within that disc. This is important to understand several approaches used in physical channel modeling.

We have exploited the eccentro-scattering physical channel model, assuming Gaussian distributed scatterers around BS and MS, to derive the probability density function (pdf) of angle of arrival of the multipath signals at BS by simple ray-tracing approach. Our model consisted of two Gaussian functions for the distribution of scatterers around BS

and MS whereas previous Gaussian scattering models considered only one Gaussian function around MS only.
The model is very useful in simulating several propagation scenarios for wireless communication systems. Equation (19) presents the formula for the pdf of AoA of the multipath signals at BS in closed form, which can be used further for beamwidth design purposes of multi antenna systems used in low antenna height environments especially in spatial division multiple access (SDMA) and wireless local area networks (WLANs). Simulated and theoretical results for such low antenna height environments have been presented which show very good proximity. Furthermore, the proposed model and the derived results are helpful in high data rate services such as wireless personal area networks (WPANs) and wireless broadband PAN (WB-PANs), which need accurate physical channel models to combat fading effects for enhanced quality of service (QoS) requirement.

References

[1] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications", *IEEE Sig. Proces. Mag.*, vol. 14, no. 6, pp. 49–83, 1997.

[2] U. Martin, J. Fuhl, I. Gaspard, M. Haardt, A. Kuchar, C. Math, A. F. Molisch, and R. Thomä, "Model scenarios for direction-selective adaptive antennas in cellular mobile communication systems – scanning the literature", *Wirel. Pers. Commun.*, vol. 11, no. 1, pp. 109–129, 1999.

[3] L. M. Correia, *Wireless Flexible Personalised Communications: COST 259, European Co-operation in Mobile Radio Research*. Chichester: Wiley, 2001.

[4] A. F. Molisch, "A generic model for MIMO wireless propagation channels in macro- and microcells", *IEEE Trans. Sig. Proces.*, vol. 52, no. 1, pp. 61–71, 2001.

[5] Z. Latinovic, A. Abdi, and Y. Bar-Ness, "A wideband space-time model for MIMO mobile fading channels", in *Proc. IEEE Wirel. Commun. Netw. Conf.*, New Orleans, USA, 2003, pp. 338–342.

[6] S. M. Seedahmed, Z. M. Hussain, and P. O'Shea, "A space-time model for mobile radio channel with hyperbolically distributed scatterers", *IEEE Anten. Wirel. Propagat. Lett.*, vol. 1, no. 12, pp. 211–214, 2002.

[7] A. Y. Olenko, K. T. Wong, and E. Hui-On Ng, "Analytically derived TOA-DOA statistics of uplink/downlink wireless multipaths arisen from scatterers on an hollow-disc around the mobile", *IEEE Anten. Wirel. Propagat. Lett.*, vol. 2, no. 22, pp. 345–348, 2003.

[8] Y. Oda, K. Tsunekawa, and M. Hata, "Geometrically based directional channel model for urban mobile communication systems", in *Proc. IEEE APWC Conf. Anten. Propagat. Wirel. Commun.*, Waltham, USA, 2000, pp. 87–90.

[9] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified channel model for mobile radio systems with smart antennas", *IEE Proc. Rad. Son. Navigat.*, vol. 145, no. 1, pp. 32–41, 1998.

[10] P. Petrus, J. H. Reed, and T. S. Rappaport, "Geometrically based statistical channel model for macrocellular mobile environments", in *Proc. IEEE GLOBECOM'96*, London, UK, 1996, pp. 1197–1201.

[11] R. B. Ertel and J. H. Reed, "Angle and time of arrival statistics for circular and elliptical scattering models", *IEEE J. Select. Areas Commun.*, vol. 17, no. 11, pp. 1829–1840, 1999.

[12] J. C. Liberti and T. S. Rappaport, "A geometrically based model for line-of-sight multipath radio channels", in *Proc. IEEE VTC'96*, Atlanta, USA, 1996, pp. 844–848.

[13] R. Janaswamy, "Angle and time of arrival statistics for the Gaussian scatter density model", *IEEE Trans. Wirel. Commun.*, vol. 1, no. 3, pp. 488–497, 2002.

[14] R. J. Piechocki, G. V. Tsoulos, and J. P. McGeehan, "Simple general formula for pdf of angle of arrival in large cell operational environments", *Electron. Lett.*, vol. 34, no. 18, pp. 1784–1785, 1998.

[15] M. P. Lotter and P. Van Rooyen, "Modeling spatial aspects of cellular CDMA/SDMA systems", *IEEE Commun. Lett.*, vol. 3, no. 5, pp. 128–131, 1999.

[16] N. M. Khan, M. T. Simsim, and P. B. Rapajic, "A generalized spatial model for all cellular environments", in *Proc. IEEE SympoTIC'04*, Bratislava, Slovakia, 2004, pp. 33–38.

[17] P. Van Rooyen, M. P. Lotter, and D. Van Wyk, *Space-Time Processing for CDMA Mobile Communications*. Boston: Kluwer, 2001.

[18] Y. Chen and V. K. Dubey, "Accuracy of geometric channel-modeling methods", *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 82–93, 2004.

[19] M. K. Simon and M. S. Alouini, *Digital Communications over Fading Channels – A Unified Approach to Performance Analysis*. New York: Wiley, 2000.

[20] A. Papoulis, Probability, *Random Variables, and Stochastic Process*. New York: McGraw-Hill, 1984.

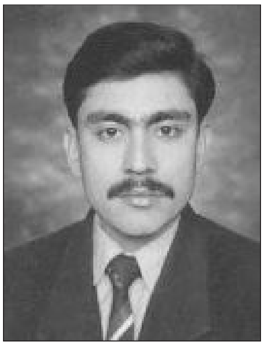
[21] C. Ward, M. Smith, A. Jeffries, D. Adams, and J. Hudson, "Characterising the radio propagation channel for smart antenna systems", *Electron. Commun. Eng. J.*, vol. 8, no. 4, pp. 191–200, 1996.

[22] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel", *IEEE J. Select. Areas Commun.*, vol. 18, no. 3, pp. 347–360, 2000.

[23] J. B. Andersen and K. I. Pedersen, "Angle of arrival statistics for low resolution antennas", *IEEE Trans. Anten. Propagat.*, vol. 50, no. 3, pp. 391–395, 2002.

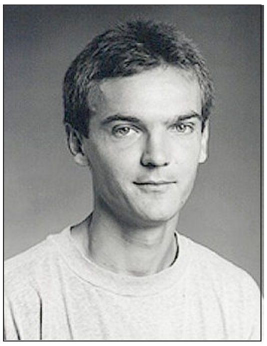
Mohammed T. Simsim was born in Makkah, Saudi Arabia, in 1975. He received the B.Sc. degree in electrical and computer engineering from UMM AL-QURA University, Makkah, Saudi Arabia, in 1998. He served in the Presidency of Civil Aviation (PCA) and UMM AL-QURA University, Saudi Arabia, from 1998 to 2001. In 2002, he received the M.Sc. degree in electrical engineering from Colorado State University (CSU), Fort Collins, CO, USA. In 2003, he joined the University of New South Wales (UNSW), Sydney, Australia, for Ph.D. studies. His research mainly focuses on physical channel modeling and its applications in cellular mobile communications.

e-mail: msimsim@student.unsw.edu.au
School of Electrical Engineering
and Telecommunications
University of New South Wales (UNSW)
Sydney, 2052, Australia



Noor M. Khan was born in Pakistan in 1973. He received the B.Sc. degree in electrical engineering from University of Engineering and Technology (UET), Lahore, Pakistan, in 1998. He held several positions in WorldCall Communications, National Institute of Science and Technical Education (NISTE), Islamabad and Pakistan Telecommunication Company Limited (PTCL) from 1998 to 2002. In 2002, he joined University of New South Wales (UNSW), Sydney, Australia, for Ph.D. studies. His research interests are channel estimation, adaptive multiuser detection, and physical channel modeling for mobile communications.

noor@unsw.edu.au
School of Electrical Engineering
and Telecommunications
University of New South Wales (UNSW)
Sydney, 2052, Australia



Predrag B. Rapajic received the B.E degree from the University of Banja Luka, Yugoslavia, in 1982, and the M.E. degree from The University of Belgrade, Belgrade, Yugoslavia, in 1988. He received the Ph.D. degree from The University of Sydney, Sydney, Australia, in 1994. In 1996, he was appointed Head of the Communications Group, Motorola Australian Research Center. Since 2005, Professor Rapajic has been working at Medway School of Engineering, University of Greenwich, UK. His research interests include adaptive multiuser detection, equalization, error control coding, mobile communication systems and multiuser information theory.

e-mail: p.rapajic@unsw.edu.au
School of Electrical Engineering
and Telecommunications
University of New South Wales (UNSW)
Sydney, 2052, Australia

Information for authors

***Journal of Telecommunications and Information Technology* (JTIT)** is published quarterly. It comprises original contributions, both regular papers and letters, dealing with a wide range of topics related to telecommunications and information technology. **All regular papers and letters are subject to peer review**. Topics presented in the JTIT report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

The JTIT is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, voice communications devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology.

Suitable research-related papers should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published only by invitation.

Manuscript. Papers published by invitation and regular papers should contain up to 15 and 8 printed pages respectively (one printed page corresponds approximately to 3 double-spaced pages of manuscript where one page contains approximately 2000 characters). An original and one copy of the manuscript should be submitted, each completed with all illustrations, tables and figure captions attached at the end of the paper. Tables and figures should be numbered consecutively with Arabic numerals. The manuscript should include an abstract about 100 words long and the relevant keywords. The abstract should contain statement of the problem, assumptions and methodology, results and conclusion or discussion on the importance of the results.
The manuscript should be double-spaced on one side of each A4 sheet (210×297 mm) only. Use of computer notation such as Fortran, Matlab, Mathematica, etc., for formulae, indices and the like is not acceptable and will result in automatic rejection of the manuscript.

Illustrations. Original illustrations should be submitted. Drawings in Corel Draw and Postscript formats are preferred. Colour illustrations are accepted only under exceptional circumstances. Lettering should be large enough to be readily legible when drawing is reduced to two- or one-column width, which often means shrinking to 1/4th of original size. Photographs should be used sparingly. All photographs should be delivered in electronic formats (TIFF, JPG, PNG or BMP). Page numbers including tables and illustrations (which should be grouped at the end) should be put in a single series, with no numbers skipped.

References. All references should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. Samples of correct formats for various types of references are presented below:

[1] Y. Namihira, "Relationship between nonlinear effective area and mode field diameter for dispersion shifted fibres", *Electron. Lett.*, vol. 30, no. 3, pp. 262–264, 1994.

[2] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1986.

[3] S. Demri and E. Orłowska,"Informational representability: Abstract models versus concrete models", in *Fuzzy Sets, Logics and Knowledge-Based Reasoning*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301–314.

Biographies and photographs of authors. A brief professional author's biography of up to 100 words and a photo of each author should be included with the manuscript. A printed photo, minimum size 35×45 mm, is acceptable. The photo may be supplied as image file.

Electronic form. TEX and LATEX are preferable, standard Microsoft Word format (.doc) is acceptable. The JTIT LATEX style file is available to authors. The file(s) should be submitted by e-mail or on a floppy disk or CD together with the hard copy of the manuscript. It is important to ensure that the diskette version and the printed version are identical. The diskette should be labelled with the following information: a) the operating system and word processing software used; b) in case of UNIX media, the method of extraction (i.e., tar) applied; c) file name(s) related to manuscript. The diskette or CD should be properly packed in order to avoid possible damage in the mail.

Galley proofs. Authors should return proofs as soon as possible. In other cases, the article will be proof-read against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within two weeks after receiving the offprint.

Copyright. Manuscript submitted to JTIT should not be published or simultaneously submitted for publication elsewhere. By submitting a manuscript, the author(s) agree to automatically transfer the copyright for their article to the publisher, if and when the article is accepted for publication. The copyright comprises the exclusive rights to reproduce and distribute the article, including reprints and all translation rights. No part of the present JTIT should not be reproduced in any form nor transmitted or translated into a machine language without prior written consent of the publisher.

A copy of the JTIT is provided to each author of paper published.

Journal of Telecommunications and Information Technology has entered into an electronic licensing relationship with EBSCO Publishing, the world's most profilic aggregator of full text journals, magazines and other sources. The full text of *Journal of Telecommunications and Information Technology* can be found on EBSCO Publishing's databases. For more information on EBSCO Publishing, please visit www.epnet.com.